

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C23

; CURRENT APPLICATION NUMBER: US/09/759,143

; CURRENT FILING DATE: 2001-01-12

; NUMBER OF SEQ ID NOS: 934

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-759-143-822

Alignment Scores:

Pred. No.: 4,07e-72 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-759-143-822 (1-675)

```
QY 1 ThrAlaLaSerAspAsnPhenGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCCCGCGTCCGATACTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGATGGCGATCGGGCGGACATCAAGCTTCCACCGCTTCATATCGGCGCTACC 141
QY 41 AlaphetLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCCGCTTGGGTGTGTGTCACACAAACGCGACGCGAGTCCAAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGAGTGCACCGCGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCGTCCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAACGGGCGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GTGACGTCTCTCGTGACCTGSCAAACCAAGTCTGGGCGGCGACGCGGTACAGGAGCGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405
```

RESULT 2

US-09-780-669-822

; Sequence 822, Application US/09780669

; Patent No. US20020051977A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, David C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedrick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C24

; CURRENT APPLICATION NUMBER: US/09/780,669

; CURRENT FILING DATE: 2001-02-09

; NUMBER OF SEQ ID NOS: 943

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-780-669-822

Alignment Scores:

Pred. No.: 4,07e-72 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-780-669-822 (1-675)

```
QY 1 ThrAlaLaSerAspAsnPhenGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCCCGCGTCCGATACTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGATGGCGATCGGGCGGACATCAAGCTTCCACCGCTTCATATCGGCGCTACC 141
QY 41 AlaphetLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCCGCTTGGGTGTGTGTCACACAAACGCGACGCGAGTCCAAACGCGTG 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGAGTGCACCGCGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCGTCCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAACGGGCGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GTGACGTCTCTCGTGACCTGSCAAACCAAGTCTGGGCGGCGACGCGGTACAGGAGCGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405
```

RESULT 3

US-09-822-827-822

; Sequence 822, Application US/09822827

; Patent No. US20020081680A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.534C1

; CURRENT APPLICATION NUMBER: US/09/822,827

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-822-827-822

Alignment Scores:

Pred. No.: 4,07e-72 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-822-827-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCAGGGATTCCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGCAGCGCATGCGATCGGGCCAGATCAAGCTCCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValAlaArgVal 60
DB 142 GCCTTCCTCGCTGGGTGTTCGACAAACAGCGACGCGATCCACCGCGTC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGACGTATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGGCGCTCCGATCAACTCCGCGATGCGCGATGCGCGATGCGCGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCACGGGTACAGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCCGAGGACCCCGGCC 405
```

## RESULT 4

US-09-895-793-822  
; Sequence 822, Application US/09895793  
; Publication No. US20020192763A1

## ; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Rual, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.534C2  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-895-793-822

## Alignment Scores:

Pred. No.: 4,07e-72 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-895-793-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGCCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCAGGGATTCCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGCAGCGCATGCGATCGGGCCAGATCAAGCTTCACCGTTTCATATCGGGCTACC 141
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValAlaArgVal 60
DB 142 GCCTTCCTCGCTGGGTGTTCGACAAACAGCGACGCGATCCACCGCGTC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGACGTATCACCGCGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
DB 262 GACGGCGCTCCGATCAACTCCGCGATGCGCGATGCGCGATGCGCGATCATCC 321
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCACGGGTACAGGAACGTG 381
QY 121 ThrLeuAlaGluGlyProProAla 128
DB 382 ACATTGGCCGAGGACCCCGGCC 405
```

## RESULT 5

US-09-895-814-822  
; Sequence 822, Application US/09895814  
; Publication No. US20020193296A1

## ; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Rual, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895,814

```
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-822

Alignment Scores:
Pred. No.: 4,07e-72 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-895-814-822 (1-675)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCGCGCTCCGATACTTCCAGCTGTCACAGGTCGGCAGGATTCGCCATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGCATGGCGATCGGGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyValAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGCTTGGTGTGTGACACACACGACGACGCGACGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGCTGATCACCGCGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaSerLeuGlyIleHisPro 100
Db 262 GACGCGCTCCGATCACTCGGCCACCGCGATGCGCGGCTTAACGGGATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 6
US-10-012-896-822
; Sequence 822, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassola, Carlota
; APPLICANT: Foy, Teresa
```

```
; APPLICANT: Fanger, Gary R.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-822

Alignment Scores:
Pred. No.: 4,07e-72 Length: 675
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-684-215B-18 (1-128) x US-10-012-896-822 (1-675)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGCGCGCTCCGATACTTCCAGCTGTCACAGGTCGGCAGGATTCGCCATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGCGCATGGCGATCGGGGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyValAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGCTTGGTGTGTGACACACACGACGACGCGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGACGCTGATCACCGCGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaSerLeuGlyIleHisPro 100
Db 262 GACGCGCTCCGATCACTCGGCCACCGCGATGCGCGGCTTAACGGGATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 7
US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
```

APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A. W.  
APPLICANT: Hepler, William T.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals y de Bassols, Carlota  
APPLICANT: Foy, Teresa M.  
APPLICANT: Katanabe, Yoshihiro  
APPLICANT: Deng, Ia  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C28  
CURRENT APPLICATION NUMBER: US/10/144,678A  
CURRENT FILING DATE: 2002-08-12  
NUMBER OF SEQ ID NOS: 1033  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-144-678A-822

Alignment Scores:  
Pred. No.: 4.07e-72 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-18 (1-128) x US-10-144-678A-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTTGGGAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGCGATGGCGATCGCGGCGCAGATCAAGCTTCCACCGTTCCATATCGGCGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlnGlyValGlnArgVal 60  
DB 142 GCCTTCCTCGCTTGGGTGTGTGTCACAAACGACGACGAGTCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGTCCGGCGCAAGTCTCGCATCTCCACCGCGACGTGATCACCGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCAACTCGGCGCACCGCATGGCGGACGGCTTAACGGGCATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 9  
US-09-736-457-1862  
Sequence 1862, Application US/09736457  
Patent No. US2002018637A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedwick, Tom  
APPLICANT: Carter, Darriek  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, Aijun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C15  
CURRENT APPLICATION NUMBER: US/09/736,457  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1862  
LENGTH: 822  
TYPE: DNA  
ORGANISM: Homo sapiens

APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A. W.  
APPLICANT: Hepler, William T.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals y de Bassols, Carlota  
APPLICANT: Foy, Teresa M.  
APPLICANT: Katanabe, Yoshihiro  
APPLICANT: Deng, Ia  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C28  
CURRENT APPLICATION NUMBER: US/10/144,678A  
CURRENT FILING DATE: 2002-08-12  
NUMBER OF SEQ ID NOS: 1033  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 822  
LENGTH: 675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-144-678A-822

Alignment Scores:  
Pred. No.: 4.07e-72 Length: 675  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-18 (1-128) x US-10-144-678A-822 (1-675)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTTGGGAGGATTCGCCATTCGGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGCGATGGCGATCGCGGCGCAGATCAAGCTTCCACCGTTCCATATCGGCGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlnGlyValGlnArgVal 60  
DB 142 GCCTTCCTCGCTTGGGTGTGTGTCACAAACGACGACGAGTCCACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGTCCGGCGCAAGTCTCGCATCTCCACCGCGACGTGATCACCGCGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGCGCTCCGATCAACTCGGCGCACCGCATGGCGGACGGCTTAACGGGCATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGTACAGGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 8  
US-10-294-025-822  
Sequence 822, Application US/10294025  
Patent No. US2003018530A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Stolk, John A.  
APPLICANT: Kalos, Michael D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C29  
CURRENT APPLICATION NUMBER: US/10/294,025



```
US-09-736-457-1862
Alignment Scores:
Pred. No.: 5.24e-72 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-736-457-1862 (1-822)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGCGTCCGATCACTCCAGCTGCCAGGGTGGGAGGATTCGGCATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValArgValGlnArgVal 60
Db 142 GCCTTCTCCGCTGGGTGGTGTTCGACAAACAGCGCACGCGCAGTCCAAACGGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGGACGTATCACCGGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCACTCGGCCACCGCATGGCGGACGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrGlnThrIlysserGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGACCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 10
US-09-902-941-1862
; Sequence 1862, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongcong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1862
Alignment Scores:
Pred. No.: 5.24e-72 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-849-626-1862 (1-822)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGCGTCCGATCACTCCAGCTGCCAGGGTGGGAGGATTCGGCATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTTACC 141
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValArgValGlnArgVal 60
Db 142 GCCTTCTCCGCTGGGTGGTGTTCGACAAACAGCGCACGCGCAGTCCAAACGGGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGGACGTATCACCGGGTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGGCGCTCCGATCACTCGGCCACCGCATGGCGGACGGCTTAACGGGCATCATCCC 321
Qy 101 GlyAspValIleSerValThrTrrGlnThrIlysserGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGACCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405

RESULT 11
US-09-849-626-1862
; Sequence 1862, Application US/09849626
; Publication No. US20020197699A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aljun
; APPLICANT: Wang, Tongcong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1862
Alignment Scores:
Pred. No.: 5.24e-72 Length: 822
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-849-626-1862 (1-822)
Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 22 ACGGCGCGTCCGATCACTCCAGCTGCCAGGGTGGGAGGATTCGGCATTCGGATC 81
```

QY 21 GlyGlnAlaMetAlaLeuGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCGACAAACGCGCAACGCGCACGCGCAGTCCAAACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 142 GCCTTCTCGCTTGGGTGTTCGACAAACGCGCAACGCGCACGCGCAGTCCAAACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACGTGATCCCGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGCTCCGATCACTCGGCACCGCATGGCGACGCGCTTAACGGGCATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 12

US-10-283-017-1862

; Sequence 1862, Application US/10283017

; Publication No. US2003021150A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Durham, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C20

; CURRENT APPLICATION NUMBER: US/10/283.017

; CURRENT FILING DATE: 2002-10-28

; NUMBER OF SEQ ID NOS: 2157

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1862

; LENGTH: 822

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-283-017-1862

Alignment Scores:

Pred. No.:	5.24e-72	Length:	822
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-283-017-1862 (1-822)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGCGCGCTCCGATACTTCAGCTGTCAGGCGGCGGCGGAGTTCGCAATCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141

QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCGACAAACGCGCAACGCGCACGCGCAGTCCAAACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 202 GTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGGACGTGATCCCGGTC 261  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
DB 262 GACGGCGCTCCGATCACTCGGCACCGCATGGCGACGCGCTTAACGGGCATCATCCC 321  
QY 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 322 GGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGCGGCACGCGTACAGGAACGTG 381  
QY 121 ThrLeuAlaGluGlyProProAla 128  
DB 382 ACATTGGCGGAGGACCCCGGCC 405

## RESULT 13

US-10-017-754-1862

; Sequence 1862, Application US/10017754

; Publication No. US20030054363A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C18

; CURRENT APPLICATION NUMBER: US/10/017.754

; CURRENT FILING DATE: 2001-10-29

; NUMBER OF SEQ ID NOS: 2004

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1862

; LENGTH: 822

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-017-754-1862

Alignment Scores:

Pred. No.:	5.24e-72	Length:	822
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-017-754-1862 (1-822)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGCGCGCTCCGATACTTCAGCTGTCAGGCGGCGGCGGAGTTCGCAATCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
DB 82 GGGCAGCGCATGGCGATCGCGGCCAGATCAAGCTTCCACCGCTTCATATCGGGCTACC 141  
QY 41 AlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArgValGlnArgVal 60  
DB 142 GCCTTCTCGCTTGGGTGTTCGACAAACGCGCAACGCGCACGCGCAGTCCAAACGCGTG 201  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80



```
; FILE REFERENCE: 210121.524C1
; CURRENT APPLICATION NUMBER: US/09/904,456
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 247
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 236
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-456-236

Alignment Scores:
Pred. No.: 5,836-72 Length: 894
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-904-456-236 (1-894)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGCGATGGCGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
QY 41 AlaphelLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCTCGGCTGGGTGTGTCGACAAACACGCGACGCGATCGGCGCTACC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTGGGAGCGCTCCGGCGGCGAGTCTCGGCATCTCCACCGCGCGATCACCGCGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
DB 262 GACGCGCTCCGATCACTCGGCGGCGAGTCTCGGCATCTCCACCGCGCGATCACCGGTC 321
QY 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGTACAGGGAACGTG 381

RESULT 17
US-09-735-705-353
; Sequence 353, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-705-353

Alignment Scores:
Pred. No.: 5,886-72 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-735-705-353 (1-900)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCCGCGTCCGATACCTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGGATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
DB 82 GGGCAGCGATGGCGATCGGGCCAGATCAAGCTTCCACCGTTTCATATCGGCGCTACC 141
QY 41 AlaphelLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
DB 142 GCCTTCTCGGCTGGGTGTGTCGACAAACACGCGACGCGATCGGCGCTACC 201
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
DB 202 GTGGGAGCGCTCCGGCGGCGAGTCTCGGCATCTCCACCGCGCGATCACCGCGGTC 261
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
DB 262 GACGCGCTCCGATCACTCGGCGGCGAGTCTCGGCATCTCCACCGCGCGATCACCGGTC 321
QY 101 GlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
DB 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGCGTACAGGGAACGTG 381

RESULT 18
US-09-850-716A-353
; Sequence 353, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-353

Alignment Scores:
Pred. No.: 5,886-72 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
```

US-09-684-215B-18 (1-128) x US-09-850-716A-353 (1-900)

```
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCGCGTCCGATACTCCAGCTGTCCAGGGTGGGAGGATTCCGATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
Db 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCAAGCTTCCACCGTTTCATATCGGGCTTACC 141
Qy 41 AlaphetLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGAACCGCGCACGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGTCCGCGGCAAGTCTCGCATCTCCACCGTTCATATCGGGCTTACC 141
Qy 41 AlaphetLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGAACCGCGCACGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGTCCGCGGCAAGTCTCGCATCTCCACCGTTCATATCGGGCTTACC 141
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCGTCCGATCACTCGGCCACCGCATGCGGAGCGGCTTAAACGGGCTATCATCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTGGTGACCTGGCAACCAAGTCCGGCGGCGACCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405
```

## RESULT 19

US-09-897-778-353

```
; Sequence 353, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-353
```

Alignment Scores:

Pred. No.:	5,88e-72	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-897-778-353 (1-900)

```
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCGCGTCCGATACTCCAGCTGTCCAGGGTGGGAGGATTCCGATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
```

```
Db 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCAAGCTTCCACCGTTCATATCGGGCTTACC 141
Qy 41 AlaphetLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
Db 142 GCCTTCTCGGCTTGGGTGTGTGTCGACAAACGCGAACCGCGCACGAGTCCACGCGTG 201
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
Db 202 GTCGGAGCGTCCGCGGCAAGTCTCGCATCTCCACCGGCGACGTATCATCGCGGCTC 261
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100
Db 262 GACGCGCGTCCGATCACTCGGCCACCGCATGCGGAGCGGCTTAAACGGGCTATCATCC 321
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
Db 322 GGTGACGTCTCTGGTGACCTGGCAACCAAGTCCGGCGGCGACCGGTACAGGGAACGTG 381
Qy 121 ThrLeuAlaGluGlyProAla 128
Db 382 ACATTGGCGGAGGACCCCGGCC 405
```

## RESULT 20

```
US-10-007-700-353
; Sequence 353, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckham, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-007-700-353
```

Alignment Scores:

Pred. No.:	5,88e-72	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-684-215B-18 (1-128) x US-10-007-700-353 (1-900)

```
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCGCGTCCGATACTCCAGCTGTCCAGGGTGGGAGGATTCCGATTCGGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
```

Db 82 GGGAGCGGATGGGATCGCGGCGCAGATCAAGTTCCCAACCGTTTCATATCGGGCTACC 141  
Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCCTCGCTGGGTGGTGTTCGACACAAACGCGCAACGGCGAGTCCCAACGCGTG 201  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCGCGCGCAAGTCTCGGCATCTCCACCGCGCAGTGCATCACCGCGTC 261  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381  
Qy 121 ThrLeuAlaGluGlyProProAla 128  
Db 382 ACATTGGCGGAGGACCCCGGCC 405  
RESULT 21  
US-10-117-982-353  
; Sequence 353, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Patrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-117-982-353  
Alignment Scores:  
Pred. No.: 5,88e-72 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0  
US-09-684-215B-18 (1-128) x US-10-117-982-353 (1-900)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGCGCGCTCCGATCAACTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGAGGCGATGGCGATCGCGGCGCAGATCAAGTTCGCCCGTTTCATATCGGGCTACC 141  
Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCCTCGGCTTGGGTGGTGTTCGACACAAACGCGCACGCGCAGGATCCAAACGCGTG 201  
US-09-684-215B-18 (1-128) x US-10-117-982-353 (1-900)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGCGCGCTCCGATCAACTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGAGGCGATGGCGATCGCGGCGCAGATCAAGTTCGCCCGTTTCATATCGGGCTACC 141  
Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCCTCGGCTTGGGTGGTGTTCGACACAAACGCGCACGCGCAGGATCCAAACGCGTG 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCGCGCGCAAGTCTCGGCATCTCCACCGCGCAGTGCATCACCGCGTC 261  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381  
Qy 121 ThrLeuAlaGluGlyProProAla 128  
Db 382 ACATTGGCGGAGGACCCCGGCC 405  
RESULT 22  
US-10-313-986-353  
; Sequence 353, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313,986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-313-986-353  
Alignment Scores:  
Pred. No.: 5,88e-72 Length: 900  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0  
US-09-684-215B-18 (1-128) x US-10-313-986-353 (1-900)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGCGCGCTCCGATCAACTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGAGGCGATGGCGATCGCGGCGCAGATCAAGTTCGCCCGTTTCATATCGGGCTACC 141  
Qy 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCCTCGGCTTGGGTGGTGTTCGACACAAACGCGCACGCGCAGGATCCAAACGCGTG 201  
US-09-684-215B-18 (1-128) x US-10-313-986-353 (1-900)  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCGGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGCAGTGCATCACCGCGTC 261  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGGCGCTCCGATCAACTCGGCACCGCGATGGCGGACGCGCTTAACGGGCATCATCC 321  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTCTCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProProAla 128  
Dd 382 ACATTGCCGAGGGACCCCGGCC 405

## RESULT 23

US-09-759-143-834  
; Sequence 834, Application US/09759143  
; Patent No. US20020022248A1

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun C.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carver, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasini A.W.

```

; APPLICANT: Hepler, William  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09/759,143  
CURRENT FILING DATE: 2001-01-12

```

; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834

```

```

; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-834

```

Alignment Scores:		
Pred. No.:	6,01e-72	Length:
Score:	653.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	100.00%	Indels:
DB:	9	Gaps:
		0
		0
		915

US-09-684-215B-18 (1-128) x US-09-759-143-834 (1-915)

QY	1	ThrAlaAlaSerAspAsnPheGlnIleuSerGlnGlyGlnGlyPheAlaIleProIle	20
pI	22	ACGGCCGGCCCGAGTAACTTCAGGTGTCCAGGGTGGCGAGGGATTCGGCATTCGGATC	81

21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40

82	GGCGAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCCCACCGTTATATCGGGCTTACC	141
D <b>b</b>		
C <b>v</b>	41 A l a b h e l e u g l v i e u c l v a l v a l a b e n d e n g l v a b e n g l v a l a b a r v i c l n a r c v a l	60

142 GCCTTCCTCGGCTGGGTGTTGTCGCAACACGGGCAACGGCGCACGAGTCCACCGTG 201

[illegible]

81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro	1000
QY		

Qy	121	ThrLeuAlaGluClyProProAla	128
Db	382	ACATTGGCCGAGGGACCCCCGGCC	405

RESULT 24

US-09-780-669-834  
; Sequence 834, Application US/09780669  
; Patent No. US20020051977A1

```

: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Devin C.
: APPLICANT: Micham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yuqi
: APPLICANT: Henderson, Robert A.
: APPLICANT: Kalos, Michael D.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stolk, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Li, Samuel
: APPLICANT: Wang, Aijun
: APPLICANT: Skelly, Vasil A.W.
: APPLICANT: Hepler, William
: APPLICANT: Hural, John
: APPLICANT: McNeill, Patricia D.

```

APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

```

; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09

```

```

; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915

```

```

: ORGANISM: Homo sapiens
: US-97-780-669-834

Alignment Scores:
  Pred No.:      6,01e-72
  Score:         653.00
  Percent Similarity: 100.00%
  Best Local Similarity: 100.00%
  Query Match:    100.00%
  DB:             9
  Length:         915
  Matches:        128
  Conservative:   0
  Mismatches:     0
  Indels:         0
  Gaps:           0

```

US-09-684-215B-18 (1-128) x US-09-780-669-834 (1-915)

QY 1 ThrAlaAlaSerAspAsnPheGlnIleuSerGlnGlyGlnGlyPheAlaIleProIle 20  
pb 22 ACGGCGCGTCCGATAACTTCAGGTGCCAGGGTGGGACGGATTCGCCATTCGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40

**D**

82  Db  Cv	GCGGACGGCATGCCATCGCGGCAGATCAAGCTTCCACCGTTTCATTATATCGGCGCTACC TTTTTTTGTGTTTGTTTGAAGAATAAACAATTAAGAAAGAACAAAAGGAAAAAAGGAAAGG AAlAbheLeucivJenUcVwAlValBanAsnBengIvaAnGIvalAcArCoValGiInArgVal	60
------------------------	---	----

142 GCCTTCCTCGGCTTGGGTGTTGCGAACCAACGGCAACGGCGCACGAGTCCACCGGTG 20

Qy 61 ValGlySerAlaProAlaAlaSerIleuGlyIleSerThrGlyAspValIleThrAlaVal 80  
rh 202 ATCCGGAGAGCACTCCGCGCGGAGAGAGTCTCCGCGATCTCCACCGGAGACGATATACGCGGTC 260

81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro	10
----	--	----

Db 322 GGTGACGTATCTCGGTGACTGGCAACCAAGTCGGGGCGGACGGGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProProAla 128

Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 25

US-09-822-827-834

; Sequence 834, Application US/09822827

; Patent No. US20020081680A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.534C1

; CURRENT APPLICATION NUMBER: US/09/822,827

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 834

; LENGTH: 915

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-822-827-834

Alignment Scores:

Pred. No.: 6,01e-72 Length: 915

Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-822-827-834 (1-915)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 22 ACGGCCGCGTCCGATACTCCAGCTGCCAGGTGGCGAGGATTCGCCATTCGCATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40

Db 82 GGGCAGCGATGGCGATCGCGGGCGACATCAAGCTTCCACCCTTCATATCGGGCTACC 141

Qy 41 AlaphelLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArgVal 60

Db 142 GCCTTCCTCGGCTTGGGTGTGTGACAAACAGGCAACGGCGCAGAGTCCACCGGTG 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

Db 202 GTCCGGAGCGTCCGGCGCAAGTCTCGCATCTCCACCGCGACGATCACCGCGGTC 261

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100

Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCGATCACC 321

Qy 101 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 120

Db 322 GGTGACGTATCTCGGTGACTGGCAACCAAGTCGGGGCGGACGGGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProProAla 128

Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 26

US-09-825-793-834

; Sequence 834, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yudiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedwick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 834

; LENGTH: 915

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-834

Alignment Scores:

Pred. No.: 6,01e-72 Length: 915

Score: 653.00 Matches: 128

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-895-793-834 (1-915)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 22 ACGGCCGCGTCCGATACTTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGCATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThr 40

Db 82 GGGCAGCGATGGCGATCGCGGGCGACATCAAGCTTCCACCCTTCATATCGGGCTACC 141

Qy 41 AlaphelLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArgValGlnArgVal 60

Db 142 GCCTTCCTCGGCTTGGGTGTGTGACAAACAGGCAACGGCGCAGAGTCCACCGGTG 201

Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

Db 202 GTCCGGAGCGTCCGGCGCAAGTCTCGCATCTCCACCGCGACGATCACCGCGGTC 261

Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100

Db 262 GACGGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGCTTAACGGGCGATCACC 321

Qy 101 GlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsnVal 120

Db 322 GGTGACGTATCTCGGTGACTGGCAACCAAGTCGGGGCGGACGGGTACAGGGAACGTG 381

Qy 121 ThrLeuAlaGluGlyProProAla 128

Db 382 ACATTGGCCGAGGACCCCGGCC 405

RESULT 27

US-09-895-814-834

; Sequence 834, Application US/09895814

; Publication No. US20020193296A1



GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriek  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Meacher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895.814  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 990  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 834  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-895-814-834

Alignment Scores:  
Pred. No.: 6,01e-72 Length: 915  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215B-18 (1-128) x US-09-895-814-834 (1-915)  
Qy 1 ThrAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTTACC 141  
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGGTGTTCGACAAACACGCAACGCGCAGCAGTCCACGCGTG 201  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCCGATCTCCACCGCGCAGTATCACC CGCGTC 261  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGCGCTCCGATCAACTCGGCCACCGGATGGCGAGCGCTTAACGGGCAATCATCCC 321  
Qy 101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 322 GGTGACGTCACTTCGGTCACTGGCAACCAAGTCCGGCGGCAACGCGTACAGGGAACGTG 381  
Qy 121 ThrLeuAlaGluGlyProAla 128  
Db 382 ACATTGGCGGAGGAGCCCCCGGCC 405

RESULT 28  
US-10-012-896-834  
; Sequence 834, Application US/10012896  
; Publication No. US20020183251A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriek  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Meacher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C27  
; CURRENT APPLICATION NUMBER: US/10/012,896  
; CURRENT FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 1011  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 834  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-012-896-834

Alignment Scores:  
Pred. No.: 6,01e-72 Length: 915  
Score: 653.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-684-215B-18 (1-128) x US-10-012-896-834 (1-915)  
Qy 1 ThrAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCAAGCTTCCACCGTTTCATATCGGGCTTACC 141  
Qy 41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 142 GCCTTCTCGGCTTGGGTGTTCGACAAACACGCAACGCGCAGTCCACGCGTG 201  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
Db 202 GTCCGGAGCGCTCCGGCGCAAGTCTCCGATCTCCACCGCGCAGTATCACC CGCGTC 261  
Qy 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPro 100  
Db 262 GACGCGCTCCGATCAACTCGGCCACCGGATGGCGAGCGCTTAACGGGCAATCATCCC 321

Qy	101	GlyAspValIleSerValThrTyrGlnThrIlysserGlyGlyThrArgThrGlyAsnVal	120
Db	322	GGTGACGTATCTCGTGACTTGGCAACCAAGTCGGCGGCACGCGTACGGGACGHS	381
Qy	121	ThrLeuAlaGlnGlyProProAla	128
Db	382	ACATGGCCGAGGACCCCGGCC	405

RESULT 29  
US-10-144-678A-834  
; Sequence 83, Application US/10144678A  
; Publication No. US20030157089A1  
; GENERAL INFORMATION:

Alignment Scores:	
Pred. No.:	6,018-72
Score:	653.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
Da:	15
Length:	915
Matches:	128
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-684-215B-18 (1-128) x US-10-144-678A-834 (1-915)

Qy	1	ThrAlaAaSerAspAsnPheGlnLeuSerGlnGlyGlycIIndlyPheAlaIleProIle	20
Db	22	ACGGCCGCTCCGATAACTTCCAGCTGTCCACGGTGGGACGGATTCGCCATTCGCATC	81
Qy	21	GlyGlnAlaMerAlaAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr	40
Db	82	GGCAGGCGATGGCGATTCGGGCCAGATCAAGCTTCCCACCGTTTCATATCGGGCTACC	141
Qy	41	AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArgValGlnArgVal	60
Db	142	GCCTTCTCGCTTGGTGTGTGTCACAAACACGGACACGGCCACGAGTCCACACGCGTG	201
Qy	61	ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal	80

Db	202	GTGGGAGCGCTCGGGCGGCAAGTCTCGGCATCTCCACGGCGACGTGATCATCCGCGTC	261
Qy	81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaaspAlaIeuAsnGlyHisHisPro	100
Db	262	GACGGCGTTCGATCAACTCGGCCACCGCATGGCGACGGCTTACGGGGATCATCCC	321
Qy	101	GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal	120
Db	322	GGTGACGTCACTCGGTGACCTGCACAAACCAAGTCGGGGCGACGGTACAGGGAACGTG	381
Qy	121	ThrIeuAlaGluGlyProProAla	128
Db	382	ACATTGGCCGAGGGACCCCGGCC	405

RESULT 30  
US-10-294-025-834  
; Sequence 834, Application US/10294025  
; Publication No. US20030185830A1  
; GENERAL INFORMATION:

Alignment Scores:	6.01e-72	Length:	915
Pred. No.:	633.00	Matches:	128
Score:		Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DR:	15		

US-09-684-215B-18 (1-128) x US-10-294-025-834 (1-915)

Qy	1	ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	22	ACGGCCGGCTCCGATAACTTCAGGTGTCCAGGGTGGCAGGATTCGCATTCGGATC	81
Qy	21	GlyGlnAlaSerAlaIleAlaGlnIleIleIysLeuProThrValHisIleGlyProThr	40
Db	82	GGCAGCGCATCGCATCGCGGGCCAGATCAAGTCTCCACCGTTCATATCGGGCTTACC	141
Qy	41	AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal	60
Db	142	GCCTTCTCTGGCTTGGGTGTGTGCACAAACACGGCAACGGCGCACGAGTCCAACGCGTG	201
Qy	61	ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal	80
Db	202	GTCCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGCAGTGATCACCGCGGTC	261
Qy	81	AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro	100
Db	262	GACGGCGCTCCGATCAACTCGGCGCACCGGATGGCGACGGGCTTAACGGGCGCATATCCC	321
Qy	101	GlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArgThrGlyAsnVal	120
Db	322	GGTGACGTATCTCGGTGACCTGGCAACCAACGAATCGGGCGGCACGCGTACAGGGAACGTG	381
Qy	121	ThrLeuAlaGluGlyProProIle	128
Db	382	ACATTGGCCGAGGACCCCGGCC	405

Search completed: April 30, 2004, 05:10:24  
Job time : 471.379 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 22:25:27 ; Search time 2507.02 Seconds  
(without alignments)  
1524.660 Million cell updates/sec

Title: US-09-684-215B-18  
Perfect score: 653  
Sequence: 1 TRAASDNFQLSQGSGQFAIRP.....QTKSGTRGTGNTLAEGPPA 128

Scoring table:  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 550265578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgm2\_1/USPTO.spool/US09684215/runat\_29042004\_061305\_13200/app\_query.fasta\_1.1180  
-DB=EST -OPT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=exn -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215@cgm2\_1\_5167@runat\_29042004\_061305\_13200 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	145.5	22.3	289	29	U82114	U82114 U82114 orde
C 2	125.5	19.2	1438	28	BH770798	BH770798 LLMT3854
C 3	123	18.8	959	28	BZ549048	BZ549048 pacsl-60
C 4	122.5	18.8	726	14	CB679186	CB679186 OSUNEF02F
C 5	119.5	18.3	590	14	CB925602	CB925602 ABA1_22_F
C 6	117.5	18.0	807	14	CA228900	CA228900 SQGFL305
C 7	114	17.5	758	13	BQ514888	BQ514888 EST622303
C 8	106	16.2	603	28	AZ934428	AZ934428 BJ_BA000
C 9	104	15.9	765	9	AJ558965	AJ558965 AJ558965
C 10	100.5	15.4	1033	28	BZ561390	BZ561390 pacsl-164
C 11	100	15.3	947	28	BZ549047	BZ549047 pacsl-60
C 12	99	15.2	726	28	AQ989479	AQ989479 RICO0025
C 13	98	15.0	1093	13	BU557763	BU557763 AGENCOURT
C 14	97	14.9	444	13	CA148171	CA148171 SCEZR2101
C 15	96.5	14.8	790	28	BZ577987	BZ577987 meh2_5667
C 16	96	14.7	546	10	AW285510	AW285510 LG1_241-E
C 17	96	14.7	551	10	AW285527	AW285527 LG1_241-G
C 18	95.5	14.6	677	12	B1378928	B1378928 BFLG1_000
C 19	95	14.5	543	12	B1721127	B1721127 1031054B1
C 20	95	14.5	673	12	B1956121	B1956121 HVSMEM002
C 21	94.5	14.5	1127	28	BZ560689	BZ560689 pacsl-164
C 22	94	14.4	982	14	CD080132	CD080132 MA3-9999U
C 23	93.5	14.3	771	29	CC322895	CC322895 OG0DK42TV
C 24	93.5	14.3	787	29	CC634156	CC634156 OG0BL3PTV
C 25	93.5	14.3	885	28	BZ567978	BZ567978 pacsl-164
C 26	92	14.1	703	14	CD383492	CD383492 PTMM08663
C 27	91.5	14.0	600	12	BG143271	BG143271 ia95h02.Y
C 28	91	13.9	701	14	CD204316	CD204316 HS1_7_H09
C 29	91	13.9	1098	13	BU553475	BU553475 AGENCOURT
C 30	90.5	13.9	631	14	CF512107	CF512107 CABUD0004
C 31	90.5	13.9	728	12	BZ285991	BZ285991 BJ285991
C 32	90	13.8	736	14	CA221763	CA221763 SCZFL404
C 33	89	13.6	509	10	BE195909	BE195909 HVSMEM009
C 34	88	13.5	655	14	CB921057	CB921057 VPD048E11
C 35	88	13.5	999	12	BM051288	BM051288 603634270
C 36	87.5	13.4	600	28	CC345457	CC345457 OGOAF70TH
C 37	87.5	13.4	643	9	AL883551	AL883551 AL883551
C 38	87.5	13.4	719	28	AZ933900	AZ933900 BJ_BA000
C 39	87.5	13.4	898	10	BQ309796	BQ309796 601892449
C 40	87	13.3	510	13	BQ977384	BQ977384 QHI24M07.
C 41	87	13.3	540	28	BZ893820	BZ893820 HL7_0146
C 42	87	13.3	642	12	BJ283688	BJ283688 BJ283688
C 43	87	13.3	643	12	BZ254998	BZ254998 BJ254998
C 44	87	13.3	649	13	BQ801199	BQ801199 WHE2811.E
C 45	87	13.3	656	13	BQ016293	BQ016293 QHE12L01.

ALIGNMENTS

RESULT 1  
U82114  
LOCUS U82114  
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
cosmid L-373; contig 64, genomic survey sequence.  
ACCESSION U82114.1 GI:3647212  
KEYWORDS GSS.  
SOURCE Mycobacterium leprae  
ORGANISM Mycobacterium leprae  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 289)

## AUTHORS

Silbaq, F.S., Cho, S.N., Cole, S.T. and Brennan, P.J.

TITLE Characterization of a 34-kilodalton protein of *Mycobacterium leprae* that is isologous to the immunodominant 34-kilodalton antigen of *Mycobacterium paratuberculosis*

JOURNAL Infect. Immun. 66 (11), 5576-5579 (1998)

MEDLINE 99003183

COMMENT 9784577

CONTACT: Silbaq FS

MICROBIOLOGY

COLORADO STATE UNIVERSITY

FORT COLLINS, CO 80523, USA

EIGLMEIER, K., HONORE, N., WOODS, S.A., CAUDRON, B. and COLE, S.T. Use

of an ordered cosmid library to deduce the genomic organization of

*Mycobacterium leprae*. Mol. Microbiol. 7 (2), 197-206 (1993)

Class: unknown.

## FEATURES

source

1. -289 Location/Qualifiers

/organism="Mycobacterium leprae"

/mol\_type="genomic DNA"

/db\_xref="taxon:1769"

/clone="cosmid L-373; contig 64"

/clone\_lib="ordered cosmid library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.19e-05 Length: 289  
 Score: 145.50 Matches: 35  
 Percent Similarity: 58.82% Conservative: 15  
 Best Local Similarity: 41.18% Mismatches: 34  
 Query Match: 22.28% Indels: 1  
 DB: 29 Gaps: 1

US-09-684-215B-18 (1-128) x U82114 (1-289)

QY 41 AlapheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---GlyAlaArgValGlnArg 59  
 Db 7 GCGTCGCGGGTGTACAGTGTGCCACCCAGCAAGCGCCGCGCGCGCAAGTATGGAC 66  
 QY 60 ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 79  
 Db 67 GTCGTGGCGGGTGGCGCGCGAATCCCGCGTCCCAAGGGGCGTGTCTCACTAAG 126  
 QY 80 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 99  
 Db 127 GTCGACGACCGCTCATCAGTAGCCGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 186  
 QY 100 ProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGlyAsn 119  
 Db 187 CCGGTGACAAAGTGTGCTGACCTATCAGGATCAGTCTGGTAGCGTCCACGGTTCAG 246  
 QY 120 ValThrLeuAlaGlu 124  
 Db 247 GTCACACTGGCAAG 261

## RESULT 2

BH770798/c

LOCUS

DEFINITION LLMtag541 MG1363 Random Sequence Tag Library Lactococcus lactis

subsp. cremoris genomic, genomic survey sequence.

ACCESSION BH770798

VERSION BH770798.1 GI:20373755

KEYWORDS GSS.

SOURCE Lactococcus lactis subsp. cremoris

ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Lactococcus.

REFERENCE 1 (bases 1 to 1438)

AUTHORS Boldrin, A., Ehrlich, S.D. and Sorokin, A.

TITLE Studies of genomes of dairy bacteria *Lactococcus lactis*

JOURNAL Sci. Aliments (2002) In press

COMMENT Contact: Sorokin A

Genétique Microbienne

INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16

Fax: 33 1 34 65 25 21

Email: sorokine@jouy.inra.fr

Best homologue in strain IL1403 is htrA (95%)

Class: shotgun

High quality sequence start: 30

High quality sequence stop: 1408.

## FEATURES

source

1. -1438 Location/Qualifiers

/organism="Lactococcus lactis subsp. cremoris"

/mol\_type="genomic DNA"

/strain="MG1363"

/sub\_species="cremoris"

/db\_xref="taxon:1359"

/clone\_lib="MG1363 Random Sequence Tag Library"

/note="Vector: pSGM2; Site 1: SmaI; Library of

chromosomal fragments of *L. lactis* strain MG1363 was

prepared by partial AluI digestion or by sonication."

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0172 Length: 1438  
 Score: 125.50 Matches: 40  
 Percent Similarity: 43.28% Conservative: 18  
 Best Local Similarity: 28.85% Mismatches: 45  
 Query Match: 19.22% Indels: 31  
 DB: 28 Gaps: 5

US-09-684-215B-18 (1-128) x BH770798 (1-1438)

QY 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla----- 27  
 Db 479 GGTATTAGGTTTGTCTATCCCATCTAATGATGTTGGTAACATCATTAATAAATCTGAACT 420  
 QY 28 ---GlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly 46  
 Db 419 GATGGTAAGATTTCACGCTCTGCC-----TTAGGTATTCTGT 384  
 QY 47 ValValAsp-----AsnAsnGlyAsn----- 53  
 Db 383 ATGTTGACTATCTCAATATCAACAATGATAGTCTCAACTGAATATTCTAGCAGC 324  
 QY 54 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGly 70  
 Db 323 GTAACGTGGAGTGGTGTCTACTCTGTTCAAGCGGGTCTTCTGCTCCACAGCTGT 264  
 QY 71 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 90  
 Db 263 CTGAAGCTGGCGATGTGATAACGAAGGTGGGAGATACCGCGGTACTTCTATCAACAGAC 204  
 QY 91 MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThr 110  
 Db 203 TTACAAAGTCTCTTTACTCACACATATTAATGATCTGTGAAGTCACTACTACTAC--- 147  
 QY 111 LysSerGlyThrArgThrGlyAsnValThrLeuAlaGlu 124  
 Db 146 CGTGATGGTAATACCCACCAATGTCAAACTCTCTAAA 105

## RESULT 3

BZ549048/c

LOCUS

DEFINITION

pacsi-60\_1644.s2 pacsi-60 *Pseudomonas aeruginosa* genomic clone

BZ549048

ACCESSION

VERSION BZ549048.1 GI:27152629

KEYWORDS GSS.

SOURCE *Pseudomonas aeruginosa*

ORGANISM *Pseudomonas aeruginosa*

Bacteria; Proteobacteria; Gammaproteobacteria; *Pseudomonadales*;

*Pseudomonadaceae*; *Pseudomonas*.

REFERENCE 1 (bases 1 to 959)

AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,

BZ549048 959 bp DNA linear GSS 17-DEC-2002

TITLE Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) in press  
CONTACT: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

FEATURES  
source  
1. 959  
Location/Qualifiers  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="1-60"  
/db\_xref="taxon:287"  
/clone="pacsl-60 1644"  
/clone\_lib="pacsl-60"  
/notes="clinical isolate 1-60 Whole genomic shotgun  
library."

Alignment Scores:  
Pred. No.: 0.0171 Length: 959  
Score: 123.00 Matches: 41  
Percent Similarity: 44.88% Conservative: 16  
Best Local Similarity: 32.28% Mismatches: 46  
Query Match: 18.84% Indels: 24  
DB: 28 Gaps: 4

US-09-684-215B-18 (1-128) x BZ549048 (1-959)  
QY 4 SerAspAsnPheGlnLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21  
Db 378 TCGCAGATCTTCACCCGCTTCGGCGCTTCATGGCTGCTCTTCCGCTTCGATCGAT 319  
QY 22 GlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThrAla 41  
Db 318 GTCCGCTGACGTGCGCGCCACCATGTTGAAG-----AAAGCGCGCAAGTCAAGT 271  
QY 42 PheLeuGlyLeuGlyValValAspAsnAsnGlyAsn----- 53  
Db 270 CGCGCTGCTGGCGGTGGTGTCCAGGAAGTGAACAGGATCTCCGCGGTCTACGGC 211  
QY 54 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 68  
Db 210 CTCGACAAGCGTCCGCGCGCTGCTGGCGCAACTGGTGAAGACGCGCGCGGCAAG 151  
QY 69 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88  
Db 150 GGTGGCTGCGAGGTGGCGGTGATCATAGCTTGAACGCGCAGTCAACGAGTCC 91  
QY 89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 108  
Db 90 GCCGACCTGCCGACCTGGTGGCAACATGAAGCGCGCGCAAGATCAACCTG----- 37  
QY 109 GlnThrLysSerGlyGlyThr 115  
Db 36 -----GACGGGGGATCC 25

# ORIGIN

Alignment Scores:  
Pred. No.: 0.0171 Length: 959  
Score: 123.00 Matches: 41  
Percent Similarity: 44.88% Conservative: 16  
Best Local Similarity: 32.28% Mismatches: 46  
Query Match: 18.84% Indels: 24  
DB: 28 Gaps: 4

US-09-684-215B-18 (1-128) x BZ549048 (1-959)  
QY 4 SerAspAsnPheGlnLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21  
Db 378 TCGCAGATCTTCACCCGCTTCGGCGCTTCATGGCTGCTCTTCCGCTTCGATCGAT 319  
QY 22 GlnAlaMetAlaIleAlaGlyGlnIleLeuProThrValHisIleGlyProThrAla 41  
Db 318 GTCCGCTGACGTGCGCGCCACCATGTTGAAG-----AAAGCGCGCAAGTCAAGT 271  
QY 42 PheLeuGlyLeuGlyValValAspAsnAsnGlyAsn----- 53  
Db 270 CGCGCTGCTGGCGGTGGTGTCCAGGAAGTGAACAGGATCTCCGCGGTCTACGGC 211  
QY 54 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 68  
Db 210 CTCGACAAGCGTCCGCGCGCTGCTGGCGCAACTGGTGAAGACGCGCGCGGCAAG 151  
QY 69 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88  
Db 150 GGTGGCTGCGAGGTGGCGGTGATCATAGCTTGAACGCGCAGTCAACGAGTCC 91  
QY 89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 108  
Db 90 GCCGACCTGCCGACCTGGTGGCAACATGAAGCGCGCGCAAGATCAACCTG----- 37  
QY 109 GlnThrLysSerGlyGlyThr 115  
Db 36 -----GACGGGGGATCC 25

RESULT 4  
CB679186/c  
LOCUS  
DEFINITION OSJNEF02F19 r OSJNEF Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEF02F19 3', mRNA sequence.  
CB679186  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CB679186 726 bp mRNA linear EST 09-APR-2003  
DEFINITION OSJNEF02F19 r OSJNEF Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEF02F19 3', mRNA sequence.  
CB679186  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

1 (bases 1 to 726)  
Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,  
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.

## TITLE

Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 02 row: F column: 19

Seq primer: gga aac agc tat gac cat g.

## FEATURES

Location/Qualifiers

1. 726  
/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNEF02F19"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSJNEF"

/note="vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; Uninfected Control"

## ORIGIN

Alignment Scores:  
Pred. No.: 0.0126 Length: 726  
Score: 122.50 Matches: 41  
Percent Similarity: 42.86% Conservative: 16  
Best Local Similarity: 30.83% Mismatches: 49  
Query Match: 18.76% Indels: 27  
DB: 14 Gaps: 4

US-09-684-215B-18 (1-128) x CB679186 (1-726)

QY 10 SerGlnGlyGlnGlyPheAlaIleProIleGlnAlaMetAlaIleAlaGlyGln 29  
Db 579 ACATCTGCTGTTGGTGTGTTTCCCATCCGTCATCCACTGTTCTGAAATAGCTCTCAG 520  
QY 30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49  
Db 519 TTA-----ATTGAGTTGGAAAAGTTCCCGTGTGCTTGAATGTGAGTTT 472  
QY 50 -----AsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 471 GCTCCAGATCCATCGCATATCAGCTTAATGTTCGACTGATCTATAATATTGAGGTT 412  
QY 61 ValGlySerAlaProAlaAlaSerLeuGly----- 70  
Db 411 CTTGGGGCAGTGTGTCGACCAAGCTGTGTTCTTCTACAGTAGGGTTTCTCTGTT 352  
QY 71 ----IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89  
Db 351 ACCATTGTTCTTGGTGATGTCATTGTTCCGTCGACGTAACCTATCAAGGCAATCT 292  
QY 90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGln 109  
Db 291 GATCTATCAGGGTTCTCGATGCTTATCGCTGTGAGCAAGGTTGAGCTTGACATCA 232  
QY 110 ThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122  
Db 231 AGA-----GGCGCTGAATCTCTGGAGGTAACCTTG 202

RESULT 5  
CB925602  
LOCUS  
DEFINITION  
ABAI\_22\_F06\_b1\_A012 Abscisis acid-treated seedlings Sorghum bicolor  
CDNA\_clone ABAI\_22\_F06\_A012 3', mRNA sequence.  
CB925602  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
EST.  
Sorghum bicolor (sorghum)  
ORGANISM  
Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Sorghum.  
Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,  
Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,  
Eastman,A. and Pratt,L.H.  
An EST database from Sorghum: ABAI-treated seedlings  
Unpublished (2003)  
Other ESTs: ABAI\_22\_F06.g1\_A012  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
the Human Genome Center, University of Tokyo Institute of Medical  
Science; plant material and RNA prepared at Texas A & M University;  
sequencing done in the Laboratory for Genomics and Bioinformatics,  
University of Georgia. Sequence ends have been trimmed to exclude  
vector and regions below Phred quality 16. Three-prime sequences  
are presented as their reverse complement and have been trimmed to  
exclude polyA.  
Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)  
POLYA=yes.  
FEATURES  
source  
1..590  
Location/Qualifiers  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="IS3620C"  
/db\_xref="taxon:4558"  
/clone="ABAI\_22\_F06\_A012"  
/lab\_host="DH10B-T1 phage-resistant B. coli"  
/clone\_lib="Abscisis acid-treated seedlings"  
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The  
library was prepared from polyA+ RNA from seedlings grown  
in hydroponic culture. After 12 days, medium was  
supplemented with 1 mM abscisis acid (ABA), while leaves  
were misted with a solution of 1 mM ABA. Roots and leaves  
were harvested after 3, 6, 12, and 24 hr and material from  
all time points was combined prior to RNA isolation.  
Double-stranded cDNA was cloned prior to RNA isolation.  
Different DraIII sites of the pME18S-FL3 vector (5-prime  
DraIII site is CACTGTGTG, 3-prime DraIII site is  
CACCATGTG). XhoI excises the cDNA insert."

Alignment Scores:  
Pred. No.: 0.0191 Length: 590  
Score: 119.50 Matches: 38  
Percent Similarity: 43.41% Conservative: 18  
Best Local Similarity: 29.46% Mismatches: 40  
Query Match: 18.30% Indels: 33  
DB: 14 Gaps: 5

US-09-684-215B-18 (1-128) x CB925602 (1-590)

Qy 15 GlyPheAlaIleProIleGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThr 34  
Db 6 GGCCTTGCTATCCATCATCACTACTTAATAATCGCTCCTCAGTTA-----53

QY 35 ValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValasp-----49  
Db 54 ATTGAGTTCGGAAGATTTCGTCGTCTGAAT---GTGACCTTCCTCCGATCCA 110  
QY 50 -----AsnAsnGlyAsnGlyAlaArgValGlnArgValGlySerAla 64  
Db 111 ATTGCATATCAGCTTAATGTTGCAACGAGCTCTTATACTTAAGGTACCTGGGGCAGT 170  
QY 65 ProAlaAlaSerLeuGlyIleSer-----Thr 73  
Db 171 GCTGCAGCAAAAGCAGGCTTGTCTCAACCGCAGGGGTTTGTGGTAATATTGTTCTG 230  
QY 74 GlyAspValIleThrAlaValaspGlyAlaProIleAsnSerAlaThrAlaMetAlasp 93  
Db 231 GGTGATATCATCTGTCAGTGGACGCAACCTGTTAAGGGCAATCTGCACCTGCTGAGG 290  
QY 94 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTpGlnThrLysSerGly 113  
Db 291 GTTCTGGATGACTATGGCGTCGAGATCAGTGACCTTGACA-----332  
QY 114 GlyThrArgThrGlyAsnValThrLeu 122  
Db 333 ---ATCCGCGGAGGCTCAGAAACCCCTT 356  
RESULT 6  
LOCUS  
DEFINITION  
CA228900/c  
CA228900.1  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Saccharum.  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
1 (bases 1 to 807)  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 058 row: G column: 06  
Seq primer: T7 Promoter Primer.  
FEATURES  
source  
1..807  
Location/Qualifiers  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCQFL3058G06"  
/lab\_host="DH10B"  
/clone\_lib="Saccharum officinarum FL3"  
/note="Organ: Base of developing inflorescence (5cm-long);  
Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Base of  
developing inflorescence (5cm-long)]. cDNA was prepared  
from polyA+ mRNA using SuperScript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"





MPI fuer Zuechtungsforschung  
Carl-von-Linne Weg 10, D-50829, Germany.  
Location/Qualifiers

source

1. .765

/organism="Antirrhinum majus"  
/mol\_type="mRNA"  
/db\_xref="taxon:4151"  
/clone="018\_1\_09\_a08"  
/tissue\_type="whole plant"  
/clone\_lib="Antirrhinum majus whole plant"

ORIGIN

Alignment Scores:  
Pred. No.: 1.21 Length: 765  
Score: 104.00 Matches: 34  
Percent Similarity: 42.24% Conservatives: 15  
Best Local Similarity: 29.31% Mismatches: 41  
Query Match: 15.93% Indels: 26  
DB: 9 Gaps: 4

US-09-684-215B-18 (1-128) x AJ558965 (1-765)

QY 10 SerGlnGlyGlnGlyPheAlaIleProIleGlnAlaMetAlaIleAlaGlyCln 29  
DB 311 ACCTGGCGGGTGTAGGATTCGATCCCTTCATCGACAGTGGTAAAGATAGTGCCTCAG 370  
QY 30 IleLysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValValAsp 49  
DB 371 TTG-----ATCAATTATGGAAGTTGTCGAGCTGGTTTAAAT---GTTGAC 415  
QY 50 -----AenAenGlyAsnGlyAlaAargValGlnArg 59  
DB 416 ATAGTCCTCGATCTATTGCAAAATCAACTCAATGTTGCAAAATGGAGCTCTAATCTTGAC 475  
QY 60 ValValGlySerAlaProAlaAlaSerLeuGly----- 70  
DB 476 GTCCCTGGAATAGTTTCAGCAGTGAAGCTGGCTGTGACCTACACACGGGGTTTCGT 535  
QY 71 -----IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88  
DB 536 GGTAAATTCGTACTTGGGGATATAATTTTGGCCGTAGATGACAAACCTGTCAGGAGTAGG 595  
QY 89 ThrAlaMetAlaAspAlaLeuAenGlyHisHisProGlyAspValIle 104  
DB 596 GCGAGCTATACAAACACTGGATGATGATTCATCGGAGCAAAAGTG 643

RESULT 10  
BZ561390 1033 bp DNA linear GSS 17-DEC-2002  
LOCUS pac82-164\_3239.y3 pac82-164 Pseudomonas aeruginosa genomic clone  
DEFINITION pac82-164\_3239, genomic survey sequence.

ACCESSION BZ561390  
VERSION BZ561390  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1033)  
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sime,E.E., Hastings,M.,  
Burns,J.L., Kaul,R. and Olsen,M.V.  
TITLE Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) In press  
CONTACT: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.  
Location/Qualifiers

FEATURES

Fax: 864 656 4293  
Email: rwing@clemons.edu  
Class: BAC ends  
High quality sequence stop: 553.

Location/Qualifiers

1. .603  
/organism="Bradyrhizobium japonicum"  
/mol\_type="genomic DNA"  
/strain="USD110"  
/db\_xref="taxon:375"  
/lab\_host="E. coli"  
/clone\_lib="B. japonicum BAC library"  
/note="Vector: pindigoS36; Site\_1: HindIII"

ORIGIN

Alignment Scores:  
Pred. No.: 0.522 Length: 603  
Score: 106.00 Matches: 35  
Percent Similarity: 41.80% Conservatives: 16  
Best Local Similarity: 28.69% Mismatches: 55  
Query Match: 16.23% Indels: 16  
DB: 28 Gaps: 4

US-09-684-215B-18 (1-128) x AZ934428 (1-603)

QY 13 GlyGlnGlyPheAlaIleProIleGlnAlaMetAlaIleAlaGlyGlnIleLys--- 31  
DB 245 GGCATCGGTCTTCGATCCCGCCCAACACCTGAGACGGTGTGTCGACGCTCAGGAC 304  
QY 32 -----LeuProThrValHisIleGlyProThrAlaPheLeuGly 44  
DB 305 AAGGTTTCGGTCAGCGCGGTGGATCGCGTCGAGATTCAGCGGTGACG----- 355  
QY 45 LeuGlyValValAspAsnAsnGly-----AenGlyAlaAargValGlnArgVal 60  
DB 356 TCGGATATCGCGCAGCAGCTCGCATGAAGACCGAAGCGCGCGCTGGTGGCGGACCG 415  
QY 61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80  
DB 416 CAGCGCAGCGTCCGCGCGCAGCGCGATCGATCGCGCGCGATGATCACCTCGGTC 475  
QY 81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAenGlyHisHisPro 100  
DB 476 AACGCGAATTCGTCAAGACGCCCGCGAGCTCGCGCCGCCACCATCGCGCGCATCGCGCCC 535  
QY 101 GlyAspValIleSerValThrTripGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
DB 536 CGTGGATCGTGAAGCTTAACGTGTGCACAG---GCCAGGACATGTCGTGACCTC 592  
QY 121 ThrLeu 122  
DB 593 ACCCTC 598

RESULT 9  
AJ558965 765 bp mRNA linear EST 12-JUN-2003  
LOCUS AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone  
DEFINITION 018\_1\_09\_a08, mRNA sequence.

ACCESSION AJ558965  
VERSION AJ558965  
KEYWORDS EST.  
SOURCE Antirrhinum majus (snapdragon)

ORGANISM Antirrhinum majus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Lamiales; Antirrhinaceae; Antirrhineae;  
Antirrhinum.

REFERENCE 1 (bases 1 to 765)  
AUTHORS Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.  
TITLE Antirrhinum EST collection  
JOURNAL Unpublished (2003)  
CONTACT: Schwarz-Sommer 2  
COMMENT Molekulare Pflanzen-genetik

Email: craymond@u.washington.edu  
Class: shotgun.  
Location/Qualifiers  
1. .947  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="1-60"  
/db\_xref="taxon:287"  
/clone\_lib="pacsl-60\_1644"  
/note="clinical isolate 1-60 Whole genomic shotgun library."  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.42 Length: 947  
Score: 100.00 Matches: 38  
Percent Similarity: 45.61% Conservative: 14  
Best Local Similarity: 33.33% Mismatches: 43  
Query Match: 15.31% Indels: 20  
DB: 28 Gaps: 3  
US-09-684-215B-18 (1-128) x BZ549047 (1-947)  
Qy 4 SerAspAsnPhenGlnLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21  
Db 374 TCCGAGTCTTCAACCGTTCGGCGCTTCATGGCGCTGTCTTCCGATTCGATCGAT 315  
Qy 22 GlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla 41  
Db 314 GTCGAGTCAAGCTGCGGCGGACCTGAGTGAAG-----AAAGCCGCGAGGTCA 267  
Qy 42 PheLeuGlyLeuGlyValValAsp----- 49  
Db 266 CGCGGCTGCTGGCGCTGGT-GATCAGGAGTGAACACGATCTCCGCGAGTCTTCGCG 208  
Qy 50 ---AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 68  
Db 207 CTGACAAGCGTTCGGCGCGCTGGTGGCGCAACTGTGGAGACGCGCGCGCGGAG 148  
Qy 69 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88  
Db 147 GGTGGCTCTGAGTGGCGGATGATCAACAGCGCTGACGCGGATCGATCAACGATCC 88  
Qy 89 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 102  
Db 87 GCGACCTGCGCGACCTGCTGGTGGAAACATGAGCCGCGGAC 46  
RESULT 12  
LOCUS BZ549047 726 bp DNA linear GSS 14-AUG-2000  
DEFINITION Photorhabdus luminescens strain W14 M13 library genomic survey sequence.  
ACCESSION AQ989479  
VERSION AQ989479.1 GI:9648073  
KEYWORDS GSS.  
SOURCE Photorhabdus luminescens  
ORGANISM Photorhabdus luminescens  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.  
REFERENCE 1 (bases 1 to 726)  
AUTHORS french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T., Daborn,P.J., Bowen,D. and Blattner,F.R.  
TITLE A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
JOURNAL 20378633  
MEDLINE 10919786  
PUBMED  
COMMENT Contact: french-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK

1. .1033  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pacsl-164\_3239"  
/clone\_lib="pacsl-164"  
/note="clinical isolate 2-164 Whole genomic shotgun library."  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.46 Length: 1033  
Score: 100.50 Matches: 42  
Percent Similarity: 42.75% Conservative: 17  
Best Local Similarity: 30.43% Mismatches: 54  
Query Match: 15.39% Indels: 25  
DB: 28 Gaps: 4  
US-09-684-215B-18 (1-128) x BZ561390 (1-1033)  
Qy 4 SerAspAsnPhenGlnLeuSerGlnGlyGln-----GlyPheAlaIleProIleGly 21  
Db 713 TCCGAGTCTTCCACCGTTCGGCGCTCAATGGCGCTGTCTTCCGATTCGATCGAT 654  
Qy 22 GlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAla 41  
Db 653 GTCGCGCTGACGCTTCGGCGGACCTTGAAGAAGCC-----GGCAAGGTCA 606  
Qy 41 laPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn----- 53  
Db 605 GTCGCGGCTGCTGGCGCTGGTGCATCCAGGAGTGAACAAGATCTCCGCGAGTCTTCG 546  
Qy 54 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaAla 68  
Db 545 GCCTGACAGCGCTCGCGCGCTGGTGGCGCAACTGTGGAGACGCGCGCGGCA 486  
Qy 68 erLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 88  
Db 495 AGCGTGGCTGAGTGGCGGATGATCCTCAGCTGAACGCGCGGATCGATCAACGAGT 426  
Qy 88 laThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal---- 106  
Db 425 CGCGCGACCTGCGGACCTGGTGGGCAACATGAACCGCGCGCAAGATCAACCTGGAGC 366  
Qy 107 -----ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120  
Db 365 TGATTCGCAACGCGCAGCGCAAGTCTCTGAGCATGGCGGTAGCGAGCTT 316  
RESULT 11  
LOCUS BZ549047/c 947 bp DNA linear GSS 17-DEC-2002  
DEFINITION pacsl-60\_1644.s1 pacsl-60 Pseudomonas aeruginosa genomic clone  
ACCESSION BZ549047  
VERSION BZ549047.1 GI:27152628  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.  
REFERENCE 1 (bases 1 to 947)  
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.  
TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library  
JOURNAL J. Bacteriol. (2002) In press  
COMMENT Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM2798 row: f column: 07  
High quality sequence stop: 299.

FEATURES  
source  
1. .1093  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6585799"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 107"  
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN  
Alignment Scores:  
Pred. No.: 8.91 Length: 1093  
Score: 98.00 Matches: 44  
Percent Similarity: 37.04% Conservative: 16  
Best Local Similarity: 27.16% Mismatches: 41  
Query Match: 15.01% Indels: 62  
DB: 13 Gaps: 8

US-09-684-215B-18 (1-128) x BU557763 (1-1093)  
Qy 26 IleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThra 41  
Db 750 ATGGCGGGGGTAAAGTATCCACGCGCGCCGCCACCGTCCAGAGGATTCG 691  
Qy 42 -----PheLeuGlyValVal 48  
Db 690 GCGCGCGGGGCGAGCTCGGCGCAAGGGGGGCACTCGAGTAGGGGGTACGGGATTCG 631  
Qy 49 AspAsnAsnGlyAsnGlyAlaArgValGln----- 58  
Db 630 ---GGCAATGCTCAGCGGCTCGCTACACCGCGTAAACACCGTCTCCGTTGGAGGATCT 574  
Qy 59 -----ArgValValGly-----SerAlaProAlaAlaSerLeuGlyIleSer 72  
Db 573 CTGCGGGAGGATTAATGGGTGCAGGGACGATGCCAACCGGGCGCGCGAGTGCGG 514  
Qy 73 ThrGlyAspValIleThrAlaValAsp-----GlyAlaProIleAsnSerAla 88  
Db 513 ACGGGGACCGCTACACCTGTTAGAGTGCACCGGGTGGGGGACCC---CGTGTCCAG 457  
Qy 89 ThrAlaMetAlaAlaAla-----LeuAsn 96  
Db 456 ACGTCCATGCCGATGCCAGGTGGCGCCAGAGGATAGAGGACAGCAGCGCCTGGAT 397  
Qy 97 GlyHis-----HisProGlyAspValIleSerVal 106  
Db 396 GGCACCATCATACCGCGGGGTTGAAGGTGTCAACATATATGGAGTCA-TCTTCTCTC 338  
Qy 107 ThrTrpGlnThrIlySerGlyThrArgThrGlyAsnValThrLeuAlaGlyPro 126  
Db 337 TGTGGGCTTGGGGTTCAGGGGAGCCTCGGTTCAGCAGCAGTGGGTGTCTCTCCGGGCCA 278  
Qy 127 ProAla 128

FEATURES  
source  
1. .726  
Location/Qualifiers  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00025"  
/clone\_lib="Photobacterium luminescens strain W14 M13 library"  
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

ORIGIN  
Alignment Scores:  
Pred. No.: 3.77 Length: 726  
Score: 99.00 Matches: 38  
Percent Similarity: 45.83% Conservative: 17  
Best Local Similarity: 31.67% Mismatches: 43  
Query Match: 15.16% Indels: 23  
DB: 28 Gaps: 5

US-09-684-215B-18 (1-128) x AQ989479 (1-726)  
Qy 6 AspPheGlnLeuSerGlnGlyGln-----GlyPheAlaIleProfile 20  
Db 179 TCATTCGATAAAGTGAAATGTTGAACGCCAGAGGTTGGGTTGCTATTCCG--- 235  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40  
Db 236 -----ACTGAATCGCTACTAAGATTATTCAGAAACTTATCCGTGACGACGAGTT 286  
Qy 41 -----AlaPheLeuGlyLeu-----GlyValValAspAsnAsnGly 52  
Db 287 ATCCGGGGTTTATCGGATACATACGCGAAGAGTGCACATATTCGCTCTTCAACGGC 346  
Qy 53 Asn-----GlyAlaArgValGlnArgValValGlySerAlaProAlaAla 67  
Db 347 AATATTACACGCTTCAAGGGCTACGAGTATTCAGTAACCTACAAACGGGCTCGCAA 406  
Qy 68 SerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSer 87  
Db 407 AAAGCCGGTATTAACTGGGTGATATCATACAGGCTCATACAAACCCGCTATTCT 466  
Qy 88 AlaThrAlaMetAlaAspAlaLeuAsnGlyHisProGlyAspValIleSerValThr 107  
Db 467 GCGGGTGAAACAATGATCAAGTGGCTGA-AATACCCCGGCGAGGTGGGTNCTGTCACT 525

RESULT 13  
BU557763/c  
LOCUS  
DEFINITION BU557763 1093 bp mRNA linear EST 16-SEP-2002  
IMAGE:6585799 5', mRNA sequence.  
ACCESSION BU557763  
VERSION BU557763.1 GI:22908059  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1093)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

TEL: (44) 1225 826621  
FAX: (44) 1225 826779  
Email: bsrf@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.  
Seq primer: M13 Forward  
Class: shotgun.  
Location/Qualifiers  
1. .726  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00025"  
/clone\_lib="Photobacterium luminescens strain W14 M13 library"  
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."



QY	83	AlaProtleAsnSerAlaThrAlaMetAlaAspAlaLeuAenGlyHISHisProGlyAsp	102
Db	335	GCGGTAGTCGGCGTGGCTGGCTGCATGCCAC-----CCAGGTCAGCACCAGTCGAT	388
QY	103	ValliserValThrTrpGlnThrLysSerGlyClyThrGrnThrGlyAsnValThrLeu	122
Db	389	ATTCAAGCATCGTC-----GTCATGCC	412
QY	123	AlaGluGlyProPro	127
Db	413	GCTGAGGTTGAACCA	427
RESULT	16		
AW285510			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-684-215B-18	(1-128)	x AW285510	(1-546)
QY	50	AsnAsnGlyAenGlyAlaArgValValGlySerAlaProAlaAlaSerLeu	69
Db	56	AATGTTCCAAACGAGCTCTTATACTTAGGTACTCTGGGGCGAGTGCTGCAGCCAAAGCA	115
QY	70	GlytIleSer-----ThrdGlyAspValIleThr	78



Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

## ORIGIN

## Alignment Scores:

Pred. No.: 6.41 Length: 543  
Score: 95.00 Matches: 30  
Percent Similarity: 40.71% Conservative: 16  
Best Local Similarity: 26.55% Mismatches: 51  
Query Match: 14.55% Indels: 16  
DB: 3

US-09-684-215B-18 (1-128) x B1721127 (1-543)

QY 10 SerGlnGlyGlnGlnGlyPheAlaIleProIleGlnAlaMetAlaIleAlaGlyGln 29  
DB 153 GCATCTCCGCGTGGGCTTCGCGTCCCATCGACAGTGTGCGCGGCTGGTAGACCAG 212  
QY 30 IleIysLeuProThrValHisIleGlyProThrAlaPheLeuGlyLeuGly----- 46  
DB 213 ATCTGACTACGCGCGGTGCTGCGCGCTGCTGGGAGTCAGCTGGCGCCGCCGACG 272  
QY 47 ValValAspAsnAsnGlyAsn---GlyAlaArgValGlnArgValValGlySerAlaPro 65  
DB 273 GTGCTCAAGCAGCTGGCGCGCGGCTGCTGTGTAGAGGTGCCCAAGGCTCCGCC 332  
QY 66 AlaAlaSerLeuGlyIleSer-----Thr 73  
DB 333 GCGGAGAGCGCGGATCAGCCACCATCGCGACCGCTTCAGCGGCTCCCTGGTGTG 392  
QY 74 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 93  
DB 393 GCGGACATCATCAGCAGCATTCAGCGCAAGCGCGTAAGAAGTAATTCGACCTGGTGAG 452  
QY 94 AlaLeuAsnGlyHisHisProGlyAspValIleSerVal 106  
DB 453 GCGTGTGTAGAGAGCGCGTGGCGACACAGTCAGAGTG 491

## RESULT 20

B1956121

## LOCUS

B1956121 673 bp mRNA linear EST 19-OCT-2001  
HVSME0025L21f Hordeum vulgare green seedling EST library  
HVCDA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA

## ACCESSION

B1956121

## VERSION

B1956121.1

## KEYWORDS

EST.

## SOURCE

Hordeum vulgare subsp. vulgare

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Hordeum.

## REFERENCE

Wing, R., Close, T.J., Klein, R., Wise, R., Chin, A., Begum, D.,

Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,

Simmons, J., Oates, R., and Main, D.

Development of a genetically and physically anchored EST resource

for barley genomics: Blumeria infected Morex (compatible) seedling

cDNA library

Unpublished (2001)

Contact: Wing RA

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hq bases = 479

Seq primer: AATTAAACCTCACTAAAGGG

High quality sequence start: 5  
High quality sequence stop: 620.

## FEATURES

source

1. 673

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/culivar="Morex"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/clone="HVSME0025L21f"

/tissue\_type="green seedling leaf"

/lab\_host="TJ121"

/clone\_lib="Hordeum vulgare green seedling EST library"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; Morex (mla) plants were greenhouse grown in the R

Wise lab at Iowa State University, Ames, IA; 7 day old

green seedlings were infected with isolate 5874 of

Blumeria graminis f. sp. hordei, and leaves were harvested

24, 48 and 72 hr post-inoculation and snap frozen (Wise).

In the TJ Close lab at the University of California,

Riverside, total RNA was prepared from each sample pool,

equal quantities of all three RNA pools were combined,

poly(A) RNA was purified from the mixture, one primary

unamplified cDNA library was made, and 1 million pfu were

in vivo excised to give pBluescript SK(-) cDNA phagemids

(Chin). Phagemids were plated and picked at the Clemson

University Genomics Institute (CUGI) (Begum, Palmer,

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA

sequencing and sequence analysis were performed at CUGI

(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).

The sequence has been trimmed to remove vector sequence

and contains a minimum of 100 bases of phred value 20 or

above. For more details on library preparation and

sequence analysis see

http://www.genome.clemson.edu/projects/barley. To order

this clone see http://www.genome.clemson.edu/orders Also

see Close TJ, Wing R, Klein, R., Wise R (2001)

Genetically and physically anchored EST resources for

barley genomics. Barley Genetics Newsletter 31:29-30.

(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

## ORIGIN

## Alignment Scores:

Pred. No.: 8.86 Length: 673  
Score: 95.00 Matches: 39  
Percent Similarity: 44.35% Conservative: 16  
Best Local Similarity: 31.45% Mismatches: 49  
Query Match: 14.55% Indels: 20  
DB: 6

US-09-684-215B-18 (1-128) x B1956121 (1-673)

QY 10 SerGlnGlyGlnGlnGlyPheAlaIleProIle-----GlyGln 22

DB 6 AATCAAGTGGAGTGTGAGGCTTCGCGATCGAGGTGTGAAACAGCGTATGCTGTCAT 65

QY 23 AlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPhe 42

DB 66 GCAAGGTGTGCTTCTTCTACTGTGACCTCAGTACATGCCCTGATGTCATCTGTGT 125

QY 43 LeuGlyLeuGlyValValAspAsnGlyAsnGly----- 54

DB 126 GGGGTGATGGAGTCGTAATAGGTTGTGATGTTTACTGGGGGGAGGATTTGAGA 185

QY 55 AlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGly 74

DB 186 GAGGCGTTCACACTTCCTATTTGGGATAGTTCCTGCTGGTCCGATTAATATTAGTGTGG 245

QY 75 AspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla 94

DB 246 AGTGTCTTGGCATCAGGAT-----CCTGTGTACACGACACTGCTTTAGCCAAG--- 296

QY 95 LeuAsnGlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGly 114  
 Db 297 ---GGTGTTCACCGATTGATGTTGCTGTAAATGG---ATCCAAGCTGGAGTT 350  
 QY 115 ThrArgThrGly 118  
 Db 351 ACTCATTTGGT 362

RESULT 21  
 BZ560689/c 1127 bp DNA linear GSS 17-DEC-2002  
 LOCUS pacs2-164\_2702.x1 pacs2-164 Pseudomonas aeruginosa genomic clone  
 DEFINITION pacs2-164\_2702, genomic survey sequence.  
 ACCESSION BZ560689  
 VERSION BZ560689.1 GI:27179214  
 KEYWORDS GSS.  
 SOURCE Pseudomonas aeruginosa  
 ORGANISM Pseudomonas aeruginosa  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.  
 REFERENCE 1 (bases 1 to 1127)  
 AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
 Burns, J.L., Kaul, R. and Olsen, M.V.  
 TITLE Whole-Genome-Sequence variation among multiple isolates of  
 Pseudomonas aeruginosa library  
 JOURNAL J. Bacteriol. (2002) In press  
 COMMENT Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.  
 Location/Qualifiers  
 1. 1127  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="2-164"  
 /db\_xref="taxon:287"  
 /clone="pacs2-164\_2702"  
 /clone\_lib="pacs2-164"  
 /note="Clinical isolate 2-164 Whole genomic shotgun  
 library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 21.8 Length: 1127  
 Score: 94.50 Matches: 36  
 Percent Similarity: 41.32% Conservative: 14  
 Best Local Similarity: 29.75% Mismatches: 46  
 Query Match: 14.47% Indels: 25  
 DB: 28 Gaps: 6

US-09-684-215B-18 (1-128) x BZ560689 (1-1127)

QY 13 GlyGlnGlyPheAlaIleProIleGly-----GlnAlaMetAlaIleAla 27  
 Db 520 GGGCAGCACCAGCGCACCAGCAGGAGGTCTCGTCGTCTTTCGCGTCGGCTCCCA 461  
 QY 28 GlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPheLeuGly 46  
 Db 460 GGGGAACGTGATCTTGATCGCGCGTACTGTCTCGTCCAGATCTCTCGCGGCTTGCC 401  
 QY 47 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValGlySerAlaProAla 66  
 Db 400 -----CACCACCGCGGCTGGTGTGTCATCGCGGCTTGGCGGTGACCGCGC 350  
 QY 67 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 86  
 Db 349 GGGCGGTAGCGGTGGCGCTTGGGATCGCTCGAAGCGGTT-----GCGGTAGTGGC 296  
 QY 87 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisPheProGlyAspValIleSerVal 106

Db 295 GTGACTGGCGTCATGGGTGAC-----CGAGGTGAGCACCAGTCGATATTCACGGATC 242  
 QY 107 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 126  
 Db 241 GTC-----GTCATGCCCGCTGAGGGTGAA 218  
 QY 127 Pro 127  
 Db 217 CCA 215

RESULT 22  
 CD080132 982 bp mRNA linear EST 14-SEP-2003  
 LOCUS MA3-9999U-M294-C03-U.B MA3-0001 Schistosoma mansoni CDNA clone  
 DEFINITION MA3-9999U-M294-C03.B, mRNA sequence.  
 ACCESSION CD080132  
 VERSION CD080132.1 GI:34631128  
 KEYWORDS EST.  
 SOURCE Schistosoma mansoni  
 ORGANISM Schistosoma mansoni  
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.  
 REFERENCE 1 (bases 1 to 982)  
 AUTHORS Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,  
 Ojopi, E.P.B., Pacuola, A.C.M., Piaza, J.P., Nishiyama, M.Y. Jr.,  
 Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,  
 Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,  
 Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,  
 Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,  
 Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,  
 Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.P.M.,  
 Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.  
 TITLE Transcriptome analysis of the acoelomate human parasite Schistosoma  
 mansoni  
 JOURNAL Nat. Genet. 35 (2), 148-157 (2003)  
 MEDLINE 22879926  
 COMMENT Contact: Dr. Sergio Verjovski-Almeida  
 Departamento de Bioquímica  
 Instituto de Química - Universidade de São Paulo  
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,  
 Brasil  
 Tel: +55-11-3091-2173  
 Fax: +55-11-3091-2186  
 Email: verj@iq.usp.br  
 This sequence was derived from the FAPESP Schistosoma mansoni EST  
 Genome Project. All sequences in the project were assembled and  
 annotated. This entry and all the assembled sequences can be seen  
 in the following URL: <http://bioinfo.iq.usp.br/schisto/>  
 Plate: MA3-9999U-M294 row: 3 column: C.

FEATURES  
 source  
 1. 982  
 /organism="Schistosoma mansoni"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:6183"  
 /clone="MA3-9999U-M294-C03.B"  
 /sex="mixed pool"  
 /dev\_stage="adult"  
 /lab\_host="Mus musculus"  
 /clone\_lib="MA3-0001"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 20 Length: 982  
 Score: 94.00 Matches: 33  
 Percent Similarity: 35.94% Conservative: 13  
 Best Local Similarity: 25.78% Mismatches: 36  
 Query Match: 14.40% Indels: 46  
 DB: 14 Gaps: 5

US-09-684-215B-18 (1-128) x CD080132 (1-982)

QY 6 AspPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIleGlnAlaMetAla 25





DB: 28 Gaps: 6

US-09-684-215B-18 (1-128) X B2567978 (1-885)

QY 9 LeuSerGlnGlyGlnGlyPheAlaIleProLeGly-----GlnAla 23  
|||:|||  
Db 570 CTTGGCGTCCAGGCACACCACCGGGCACCCAGCAGGAGGAGGTCTCGTGCTTTGCC 511  
|||:|||  
QY 24 MetAlaIleAlaGlyGlnIleLysLeu---ProThrValHisIleGlyProThrAlaPhe 42  
:::|||:  
Db 510 GTGCGGGTCCAGGGGAACCTGGATCTTGATCCGGCGTACTGTCGGTCCAGATCTCCTC 451  
:::|||:  
QY 43 LeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValgly 62  
:::::  
Db 450 GCCGCCCTTCCC-----CAACCACAGTCCGTCTGGGTGTGATGCGCGGCTTGGG 400  
:::::  
QY 63 SerAlaProAlaAlaSerLeuGlyLieserThrGlyAspValIleThrAlaValAspgly 82  
:::::  
Db 399 CGTAGCCCGCGGGGCTAGGCGGTGGCGCTTGGGGATCGGCTCGAACCGGTT----- 346  
:::::  
QY 83 AlaProfileAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisProGlyasp 102  
:::|||:  
Db 345 GCGGTAGTGCCTGACTGCTGGCGTCTATGGGTGAC-----CGAGGTGAGCACCAGTCGAT 292  
:::|||:  
QY 103 ValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 122  
:::|||:  
Db 291 ATTACAGCAATCGTC-----GTGATGCC 268  
|||:|||  
QY 123 AlaGluGlyProPro 127  
|||:|||  
Db 267 GCTGAGGTGAACCA 253  
|||:|||

RESULT 26  
CD383492  
LOCUS  
DEFINITION PTTM08663 Phaeodactylum tricornutum Uni-zap XR Phaeodactylum  
tricornutum cDNA 5', mRNA sequence.  
CD383492 CD383492.1 GI:31259106  
EST.  
Phaeodactylum tricornutum  
Phaeodactylum tricornutum  
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;  
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.  
REFERENCE 1 (bases 1 to 703)  
AUTHORS Scala,S., Carls,N., Falcatore,A., Chiusano,M.L. and Bowler,C.  
TITLE Genome properties of the diatom Phaeodactylum tricornutum  
JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)  
MEDLINE 2211123  
PubMed 12114555  
COMMENT Contact: Bowler C  
Laboratory of Molecular Plant Biology  
Stazione Zoologica 'Anton Dohrn'  
Villa Comunale, I-80121, Napoli, Italy  
Tel: 39 081 583 3268/3211  
Fax: 39 081 764 1355  
Email: chris@alpha.szn.it  
Diatom EST Database(<http://aves.thagen.sznbowler.com>)  
Seq primer: T3 backward  
POLYA-No. Location/Qualifiers  
1..703  
/organism="Phaeodactylum tricornutum"  
/mol\_type="mRNA"  
/db\_xref="taxon:2850"  
/cell\_line="COMP632"  
/clone\_lib="Phaeodactylum tricornutum Uni-Zap XR"  
/notes="Vector: Uni-Zap XR vector; Site\_1: Eco RI; Site\_2:  
Xho I"

ORIGIN  
Alignment Scores: 19.6 Length: 703  
Pred. No.: 1

Score:	92.00	Matches:	35
Percent Similarity:	40.00%	Conservative:	17
Best Local Similarity:	26.92%	Mismatches:	54
Query Match:	14.09%	Indels:	25
DB:	14	Gaps:	4
US-09-684-215B-18 (1-128) x CD383492 (1-703)			
QY	10	SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla-----	27
DB	178	GCCTCGCGCGGATTGGTTTGCATCCCAATGACACCGCGAAATTTATGTTCGAGACG	237
QY	28	-----GlyGlnIleLysLeuProThrValHisIleGlyProThrAlaPhe--	42
DB	238	CTCATTCGGGATGGCAGGGTGGTTCGACCACTTTGGTATTTCCTACCTCGAATACAA	297
QY	43	-----LeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal	60
DB	298	CAAGCAGCGCGTCTTAGGATA-----AATAGCGGTGCTTGGTACTCGAAGCA	345
QY	61	ValGlySerAlaProAlaAlaSerLeuGly-----	70
DB	346	CCAGCGCGGAAGCCCTCTCTCCCGGTCTCAAAGGAACGCGACGACCGAGTCGGT	405
QY	71	---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr	89
DB	406	TTGGTAGAATCGGTGACATTATACGAAGATGGGAGACAAGTGCATACCGTGGAGTCC	465
QY	90	AlaMetAlaAspAlaLeuAsnGlyHisProGlyAspValIleSerValThrTrpGln	109
DB	466	GATTTCTTTCAAGCTTTAGACAGTACAAACCGCGGTGATGTCGTGGAGTCGACCGT-AAA	524
QY	110	ThrLysSerGlyGlyThrArgThrGlyAsn	119
DB	525	TCGGGTTTCGGCAGTGAACGATCAGCTAAC	554
RESULT 27			
EG143271/c			
LOCUS			
DEFINITION			
<p>EG143271 600 bp mRNA linear EST 13-MAR-2002</p> <p>la9Sh02.v1 Melton Mouse E16 5 Pancreas Library 2 M16B2 Mus musculus</p> <p>cDNA clone IMAGE:5646194 5', similar to SW:PM1_HUMAN P17152 PUTATIVE</p> <p>RECEPTOR PROTEIN. ;, mRNA sequence.</p>			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
<p>Bowers, Y.</p> <p>WashU-Harvard Pancreas EST Project</p> <p>Unpublished (2000)</p> <p>Contact: Douglas Melton, Klaus H. Kaestner, &amp; Hiroshi Inoue</p> <p>Endocrine Pancreas Consortium</p> <p>Harvard University, Howard Hughes Medical Institute</p> <p>Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,</p> <p>MA 02138</p> <p>Tel: 617-495-1812</p> <p>Fax: 617-495-8557</p> <p>Email: dmelton@biohpc.harvard.edu</p> <p>Library was constructed by Dr. Douglas Melton DNA sequencing by:</p> <p>Washington University Genome Sequencing Center For information on</p> <p>obtaining a clone please contact: Juliana Brown</p> <p>(brown@fas.harvard.edu)</p> <p>MEI:1852914 This sequence now available from the IMAGE consortium,</p> <p>for clone orders contact: info@image.lln.gov</p> <p>Seq primer: -40RP from Gibco</p> <p>High quality sequence stub: 455.</p>			
TITLE			
JOURNAL			
COMMENT			

CD204316	CD204316	701 bp	linear	EST 20-MAY-2003
LOCUS	HS1_7_H09_b1_A012	Heat-shocked seedlings	Sorghum bicolor	cDNA clone
DEFINITION	HS1_7_H09_A012 3'	mRNA sequence.		
ACCESSION	CD204316			
VERSION	CD204316.1	GI:30933230		
KEYWORDS	EST.			
SOURCE	Sorghum bicolor (sorghum)			
ORGANISM	Sorghum bicolor			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
	clade; Panicoideae; Andropogoneae; Sorghum.			
REFERENCE	1 (bases 1 to 701)			
AUTHORS	Cordonnier-pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,			

Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D., Eastman,A. and Pratt,D.H.  
An EST database from Sorghum: heat-shocked seedlings  
Unpublished (2003)  
Other ESTs: HS1.7.H09.G1.A012  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

## TITLE

## JOURNAL

## COMMENT

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)  
POLYA=Yes.

## FEATURES

source

1. 701

/organism="Sorghum bicolor"

/mol\_type="mRNA"

/cultivar="IS3620C"

/db\_xref="taxon:4558"

/clone="HS1.7.H09.A012"

/lab\_host="DH10B-T1 phage-resistant E. coli"

/notes="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from 6-day-old seedlings grown in hydroponic culture and heat-shocked at 40-42 C for 4 or 24 hr. After heat shock, roots and leaves were harvested and tissues combined for RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG)."

## ORIGIN

Alignment Scores:

Pred. No.:	24.9	Length:	701
Score:	91.00	Matches:	32
Percent Similarity:	38.71%	Conservative:	16
Best Local Similarity:	25.81%	Mismatches:	50
Query Match:	13.94%	Indels:	26
DB:	14	Gaps:	4

US-09-684-215B-18 (1-128) x CD204316 (1-701)

QY	16	PheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrVal	35
DB	3	TTTGCAGTACCGATTGACTCTGTCATTAATAATAGTAGAGAACTTTAAGAAAAATGGAGA	62
QY	36	HistLeGlyProThrAlaPheLeuGlyValValValAspAsn	51
DB	63	GTGTGTAAGGCCA-----TGCCTTGTTTAAAGATGCTTGACCTGAATCCATGATCAT	116
QY	52	-----GlyAsnGlyAlaArgValGln	58
DB	117	GCACAGCTCAAGAAAAAATCAAGTACTTTCCAGATGTAAGAAAGGGTGTGTCTCT	176
QY	59	ArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThr	78
DB	177	ATGGTTACACACAGATCTCAGCTGAACAAAGCAGGATTCCTCCTGGGATGTTGTT	236
QY	79	AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis	98
DB	237	GAATTTGGTGCAACCAAGTGTAGAGCATCAAGAGATCATTCATATCATG-----	287
QY	99	HisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArgThrGly	118

DB	288	-----GGGGACAAGTCGGAGTACCATTTAAAGTCTTGTGTC-----AAAAGACGAGC	335
QY	119	AsnValThrLeu 122	
DB	336	AATGTGACAGTG 347	
RESULT 29			
BUS53475			
LOCUS			
DEFINITION	BUS53475	1098 bp	mRNA
	AGENCOURT 10242287 NIH MGC 109 Homo sapiens	linear	EST 16-SEP-2002
	IMAGE:6577858 5', mRNA sequence.		
ACCESSION	BUS53475		
VERSION	BUS53475.1	GI:22903747	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1. (bases 1 to 1098)		
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC		
FEATURES	Location/Qualifiers		
source	1..1098		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6577858"		
	/tissue type="teratocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_109"		
	/note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		

## ORIGIN

Alignment Scores:

Pred. No.:	49	Length:	1098
Score:	91.00	Matches:	40
Percent Similarity:	39.85%	Conservative:	13
Best Local Similarity:	30.08%	Mismatches:	38
Query Match:	13.94%	Indels:	42
DB:	13	Gaps:	8

US-09-684-215B-18 (1-128) x BUS53475 (1-1098)

QY	23	AlaValAlaIleAlaGly-----ClnIleLysLeuProThrValHisIle	37
DB	340	GCTGCCCGGTGGAGGGCCCATACCCAGTGGAGGATCCGTCTGCCAAGAGCTGGG---	396
QY	38	GlyProThrAlaPheLeuGlyValValValAspAsnGlyAsn-----	53
DB	397	GGCCCT-----CTGGGCTTACTATTGTCGGAGCTCCGACCATCCAGCCACCCG	447
QY	54	-----GlyAlaArgValGlnArgValGlySerAlaProAlaAla	67

Db 448 TTGGTGTCCAGGACCTGGCTGTTATCTCCAGGTGCTCCCGCGGGCGCTGGCGCT 507  
Qy 68 SerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSer 87  
Db 508 CCACGCGCTCGCGGTTGGGACCGCATCTCTGGCGCTGAACGGGCAAGACGTGCGGGAT 567  
Qy 88 AlaThr-----AlaMetAlaAspAlaLeu----- 95  
Db 568 GCCACCCACCAAGAACAGTCAAGTGCCTGCTCCGCGCTGCGCTGGAGCTTGTGCT 627  
Qy 96 -----AsnGlyHisProGlyAspValIleSerValThrTrpGln 109  
Db 628 GCTGGGTGCGGAGGGACCGCGGCACCCCGCGGCGC---TACGGGCAAAACCTGGTGC 684  
Qy 110 ThrLysSer-----GlyGlyThrArgThrGly 118  
Db 685 ACTCCCCCAAAAGGACCTTGTGGCGGAACAAAGGCTGGG 723

RESULT 30  
CF512107/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

CF512107 631 bp mRNA linear EST 09-SEP-2003  
CABud0004 IVR\_E10 Vitis vinifera cv. cabernet sauvignon (Clone 8)  
Bud - CABUD Vitis vinifera cDNA clone CABud0004\_IVR\_E10 3', mRNA  
sequence.  
CF512107  
CF512107.1 GI:34543875  
EST.  
Vitis vinifera  
Vitis vinifera  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; Vitaceae; Vitis.  
1 (bases 1 to 631)  
Goes da Silva,F., Iandolo,A., Lim,H., Baek,J., Jones,K. and  
Cook,D.

Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'  
berries at various developmental stages  
Unpublished (2003)  
Contact: Douglas Cook, PhD  
CAES Genome Facility  
UC Davis, Plant Pathology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 754 6561  
Fax: 530 754 6617  
Email: drcook@ucdavis.edu  
Seq primer: GCCAACGAATGGTCTAG.  
Location/Qualifiers  
1..631  
/organism="Vitis vinifera"  
/mol\_type="mRNA"  
/cultivar="Cabernet Sauvignon (Clone 8)"  
/db\_xref="taxon:29760"  
/clone="CABud0004\_IVR\_E10"  
/dev\_stage="pre-bloom (10-11 days before bloom)"  
/lab\_host="DH5alpha"  
/clone\_lib="Vitis vinifera cv. cabernet sauvignon (Clone  
8) Bud - CABUD"  
/notes="Organ: Bud; vector: pDNR; Site\_1: Sfil; Site\_2:  
Sfil; CABUD is a cDNA library of Vitis vinifera cv.  
'Cabernet Sauvignon' Clone 8 dissected buds. Samples were  
collected May 13, 2002 from pre-bloom plants (10-11 days  
before bloom), pre-veraison. Sampled vines were located at  
the University of California, Davis, Experimental  
Vineyard. cDNAs were made by oligo-dT priming and  
directionally cloned. 5' and 3' adaptors were used in  
cloning as follows:  
5'-AAGCAGTGTATCAGCAGAGTGGCCATACGGCGGG-3' and  
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)NN-3'. Library was  
constructed using the Clontech Creator SMART kit and  
size-selected to contain the 0.5-3 kb size fraction."

FEATURES  
source

ORIGIN

Alignment Scores: 23.9 Length: 631  
Prec. No.: 90.50 Matches: 31  
Score: 44.44% Conservative: 21  
Percent Similarity: 26.50% Mismatches: 40  
Best Local Similarity: 13.86% Indels: 25  
Query Match: 14 Gaps: 5  
DB: 14

US-09-684-215B-18 (1-128) x CF512107 (1-631)

Qy 10 SerGlnGlyGlyClnGlyPheAlaIlePheAlaIleGlnAlaIleAlaGlyGln 29  
Db 574 GCATCTTCTGTGTGCGGATTTCAATTCAGTTGACACTGTAAAGCGGCATTTCGACCAA 515  
Qy 30 Ile-----LysLeuProThrValHisIleGly-----ProThrAla 41  
Db 514 TTGGTGAGGTTTGGAAAGTCCAGACCTATTTAGGAATTAAGTTTGGCGCTGATCAG 455  
Qy 42 PheLeuGly---LeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60  
Db 454 TCTGTGGGCGAGTTGGGTGTA-----AGTGGGCGCCCTTTGTTTATGATGCT 410  
Qy 61 ValGlySerAlaProAlaAlaSerLeuGly----- 70  
Db 409 CCTGCAACGCTCCTCTGCGCAAGCGGCGCTACTACCAACCAAGCTGATGCCTATGGC 350  
Qy 71 ---IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThr 89  
Db 349 AGACTTATTTAGGTGACATCATACATCGGTGAATGGAATAAGGTTTCCCAATGAAGC 290  
Qy 90 AlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 106  
Db 289 GACTTGTACAGAATCTTGCACAGTAAAGTGGGTGACACCGTGACTGTG 239

Search completed: April 30, 2004, 04:40:59  
Job time : 2513.02 secs

RESULT 30  
AX832594  
LOCUS AX832594 1872 bp DNA linear PAT 12-DEC-2003  
DEFINITION Sequence 17 from Patent EPI347055.  
ACCESSION AX832594  
VERSION AX832594.1 GI:39840644  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.  
REFERENCE  
AUTHORS Reed, S.G., Skeiky, Y.A., Campos-Neto, A., Houghton, R.L.,  
Vedvick, T.S., Twardzik, D.R. and Dillon, D.C.  
TITLE Compounds for immunotherapy and diagnosis of tuberculosis  
JOURNAL Patent: EP 1347055-A 17 24-SEP-2003;  
CORIXA CORPORATION (US)  
FEATURES  
source location/Qualifiers  
1..1872  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,338-41 Length: 1872  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservative: 0  
Best Local Similarity: 99.22% Mismatches: 1  
Query Match: 99.23% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215B-23 (1-128) x AX832594 (1-1872)  
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 758 ACGGCCGGTCCGATACATTCAGCTGCCAGGGTGGGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 818 GGGCAGCGGATGGCGATCCGGGCCCAATCCGATCGGTGGGGGTCACCACCGTTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArg 60  
DB 878 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACACACGCGCACGGCGCAGCA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 938 GTCCAAACGGGTGGTGGAGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 998 ATCACCGCGGTGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTAAC 1057  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 1058 GGGCATCATCCCGGTGACGTATCTCGGTGAATCGCAACCAAGTCGGCGGCACGGGT 1117  
QY 121 ThrGlyAsnValThrLeuAlaGlu 128  
DB 1118 ACAGGAGACGTGACATGGCCGAG 1141

Search completed: April 30, 2004, 02:25:07  
Job time : 3732.05 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 22:43:07 ; Search time 469.129 Seconds

(without alignments)  
1234.209 Million cell updates/sec

Title: US-09-684-215b-18

Perfect score: 653  
Sequence: 1 TRASNDFQLSQGGQGAIF.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Ygapext 0.5	
Ygapop 6.0 , Xgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09684215/runat\_29042004\_061306/app\_query.fasta\_1.1180  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnnp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=30 -MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09684215 @CGN 1 1 370 @runat\_29042004\_061306\_13266  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18:	/cgn2_6/ptodata/1/pubpna/US00_NEW_PUB.seq:*
19:	/cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

## ALIGNMENTS

### RESULT 1

US-09-759-143-822

; Sequence 822, Application US/09759143

; Patent No. US2002002248A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Ranger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

1	653	100.0	675	9	US-09-759-143-822	Sequence 822, App
2	653	100.0	675	9	US-09-780-669-822	Sequence 822, App
3	653	100.0	675	9	US-09-822-827-822	Sequence 822, App
4	653	100.0	675	9	US-09-895-793-822	Sequence 822, App
5	653	100.0	675	9	US-09-895-814-822	Sequence 822, App
6	653	100.0	675	14	US-10-012-896-822	Sequence 822, App
7	653	100.0	675	15	US-10-144-678A-822	Sequence 822, App
8	653	100.0	675	15	US-10-294-025-822	Sequence 822, App
9	653	100.0	822	9	US-09-736-457-1862	Sequence 1862, Ap
10	653	100.0	822	9	US-09-902-941-1862	Sequence 1862, Ap
11	653	100.0	822	13	US-10-849-626-1862	Sequence 1862, Ap
12	653	100.0	822	13	US-10-283-017-1862	Sequence 1862, Ap
13	653	100.0	822	15	US-10-017-754-1862	Sequence 1862, Ap
14	653	100.0	822	15	US-10-113-872-1862	Sequence 1862, Ap
15	653	100.0	894	13	US-09-878-722-236	Sequence 236, App
16	653	100.0	894	13	US-09-904-456-236	Sequence 236, App
17	653	100.0	900	9	US-09-735-705-353	Sequence 353, App
18	653	100.0	900	9	US-09-850-716A-353	Sequence 353, App
19	653	100.0	900	9	US-09-897-778-353	Sequence 353, App
20	653	100.0	900	13	US-10-007-700-353	Sequence 353, App
21	653	100.0	900	15	US-10-117-982-353	Sequence 353, App
22	653	100.0	900	16	US-10-313-986-353	Sequence 353, App
23	653	100.0	915	9	US-09-759-143-834	Sequence 834, App
24	653	100.0	915	9	US-09-780-669-834	Sequence 834, App
25	653	100.0	915	9	US-09-822-827-834	Sequence 834, App
26	653	100.0	915	9	US-09-895-793-834	Sequence 834, App
27	653	100.0	915	9	US-09-895-814-834	Sequence 834, App
28	653	100.0	915	14	US-10-012-896-834	Sequence 834, App
29	653	100.0	915	15	US-10-144-678A-834	Sequence 834, App
30	653	100.0	915	15	US-10-294-025-834	Sequence 834, App
31	653	100.0	945	9	US-09-736-457-1861	Sequence 1861, Ap
32	653	100.0	945	9	US-09-902-941-1861	Sequence 1861, Ap
33	653	100.0	945	9	US-09-849-626-1861	Sequence 1861, Ap
34	653	100.0	945	13	US-10-283-017-1861	Sequence 1861, Ap
35	653	100.0	945	15	US-10-017-754-1861	Sequence 1861, Ap
36	653	100.0	945	15	US-10-113-872-1861	Sequence 1861, Ap
37	653	100.0	1012	9	US-09-735-705-351	Sequence 351, App
38	653	100.0	1012	9	US-09-850-716A-351	Sequence 351, App
39	653	100.0	1012	9	US-09-897-778-351	Sequence 351, App
40	653	100.0	1012	13	US-10-007-700-351	Sequence 351, App
41	653	100.0	1012	15	US-10-117-982-351	Sequence 351, App
42	653	100.0	1012	16	US-10-313-986-351	Sequence 351, App
43	653	100.0	1035	9	US-09-928-217-1084	Sequence 1084, Ap
44	653	100.0	1035	9	US-09-838-263-1084	Sequence 1084, Ap
45	653	100.0	1035	10	US-09-938-864-388	Sequence 388, App

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2004, 23:03:22 ; Search time 291 Seconds  
(without alignments)  
5781.053 Million cell updates/sec

Title: US-09-684-215B-3

Perfect score: 396

Sequence: 1 acggcgctccgataactt.....tgccgagggaccccgcc 396

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 337363 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04.\*

- 1: geneseq1980s.\*
- 2: geneseq1990s.\*
- 3: geneseq2000s.\*
- 4: geneseq2001as.\*
- 5: geneseq2001bs.\*
- 6: geneseq2002s.\*
- 7: geneseq2003as.\*
- 8: geneseq2003bs.\*
- 9: geneseq2003cs.\*
- 10: geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	100.0	396	4	AAL40769 Nucleotid
2	396	100.0	672	4	AAL40772 Nucleotid
3	396	100.0	702	2	Aaz20206 Mycobacte
4	396	100.0	702	4	AAL40770 Nucleotid
5	396	100.0	702	6	ABK14140 DNA encod
6	396	100.0	1002	6	Aad47078 Mycobacte
7	396	100.0	1002	6	Aad47077 Mycobacte
8	396	100.0	1002	6	Aad28336 Mycobacte
9	396	100.0	1002	6	Aad28337 Mycobacte
10	396	100.0	1068	2	Aax34251 Mycobacte
11	396	100.0	1143	2	Aax34252 Mycobacte
12	396	100.0	1742	4	AAL40771 Nucleotid
13	396	100.0	2181	8	ADA26353 Mycobacte
14	396	100.0	2190	6	Aad47084 Mycobacte
15	396	100.0	2190	6	Aad28343 Mycobacte
16	396	100.0	2191	4	AAL40773 Nucleotid
17	396	100.0	2286	6	ABK14128 DNA encod
18	396	100.0	2287	2	Aaz20194 Mycobacte
19	396	100.0	2287	6	Aad47083 Mycobacte
20	396	100.0	2287	6	Aad28342 Mycobacte
21	396	100.0	2451	8	ADA26360 Mycobacte
22	396	100.0	2487	8	ADA26359 Mycobacte
23	396	100.0	2637	8	ADA26358 Mycobacte

24	396	100.0	2808	6	AAD47110
25	396	100.0	2808	8	ADA26357 Mycobacte
26	396	100.0	3030	8	ADA26355 Mycobacte
27	396	100.0	3060	8	ADA26363 M. bovis
28	396	100.0	3104	8	ADA26362 Mycobacte
29	396	100.0	3474	8	ADA26361 Mycobacte
30	396	100.0	110000	4	Continuation (2 of
31	396	100.0	110000	4	Continuation (2 of
32	394.4	99.6	447	2	AAT91403
33	394.4	99.6	447	2	AAT91466
34	394.4	99.6	447	2	AAT91466
35	394.4	99.6	447	2	AAV64450 M. tuberc
36	394.4	99.6	447	2	AAV64450 M. tuberc
37	394.4	99.6	447	2	AAZ19040 M. tuberc
38	394.4	99.6	447	5	AAZ19252 M. tuberc
39	394.4	99.6	447	5	AAZ19252 M. tuberc
40	394.4	99.6	447	6	AAZ19252 M. tuberc
41	394.4	99.6	1629	5	AAZ19252 M. tuberc
42	392.8	99.2	1871	4	AAZ19252 M. tuberc
43	391.2	98.8	1872	2	AAZ191414
44	391.2	98.8	1872	2	AAZ191414
45	391.2	98.8	1872	2	AAV44355 Mycobacte

#### ALIGNMENTS

#### RESULT 1

AL40769  
ID AAL40769 standard; DNA; 396 BP.

XX AAL40769;

XX 03-OCT-2002 (first entry)

XX Nucleotide sequence encoding Ral2 protein.

XX Ral2; serine protease antigen; Mtb32A; Mycobacterium tuberculosis;

XX vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;

XX immunogen; cytokine; gene; ds.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 1..396

XX FT /\*tag= a

XX FT /product= "Ral2 protein"

XX FT /note= "No start or stop codon"

XX PN WO200125401-A2.

XX 12-APR-2001.

XX 06-OCT-2000; 2000WO-US027652.

XX 07-OCT-1999; 99US-0158585P.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Guderian J;

XX WPI; 2001-266299/27.

XX P-PSDB; AAO22138.

XX Recombinant nucleic acid molecule for producing high yield expression of

XX desired fusion polypeptides, encodes fusion polypeptide comprising

XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide.

XX Claim 1; Fig 2; 39pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a

XX fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14

XX kDa C-terminal fragment of serine protease antigen Mtb32A of



CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2 protein  
 XX  
 XX Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 4; Length 396;

Best Local Similarity 100.0%; Pred. No. 5.2e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATTAATTCAGCTGCCAGGTCGCGAGGATTCGCCATTCGATC 60  
 Db 1 ACGGCGGCTCCGATTAATTCAGCTGCCAGGTCGCGAGGATTCGCCATTCGATC 60  
 QY 61 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGGTGGGGTCAACCCCGTTTCA 120  
 Db 61 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGGTGGGGTCAACCCCGTTTCA 120  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGTCTCGACACACGCGACGCGACGA 180  
 Db 121 ATCGGCGCTACCGCTTCTCGGCTTGGTCTCGACACACGCGACGCGACGA 180  
 QY 181 GTCCAAACGCGTGTGGAGGCGTCCCGCGCAAGTCTCGGCACTCTCCACCGCGAGTG 240  
 Db 181 GTCCAAACGCGTGTGGAGGCGTCCCGCGCAAGTCTCGGCACTCTCCACCGCGAGTG 240  
 QY 241 ATCACCGGCTCGACGCGCTCCGATCACTCGGCCACCGCGATGGCGGCGCTTAAC 300  
 Db 241 ATCACCGGCTCGACGCGCTCCGATCACTCGGCCACCGCGATGGCGGCGCTTAAC 300  
 QY 301 GGGCATATCCCGGTAGCTATCTCGTGACCTGGCAACCAAGTCGGCGGCGCGCT 360  
 Db 301 GGGCATATCCCGGTAGCTATCTCGTGACCTGGCAACCAAGTCGGCGGCGCGCT 360  
 QY 361 ACAGGGAACGTCACATTGGCGAGGACCCCGGCC 396  
 Db 361 ACAGGGAACGTCACATTGGCGAGGACCCCGGCC 396

## RESULT 2

AA140772

ID AAL40772 standard; DNA; 672 BP.

XX AC AAL40772;

XX DE 06-AUG-2003 (revised)

XX DT 03-OCT-2002 (first entry)

XX DT Nucleotide sequence encoding Ral2-mammaglobin fusion protein.

XX DE Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;

XX KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;

XX KW immunogen; cytokine; gene; ds.

XX XX Mammalia.

OS Unidentified.

OS Chimeric.

OS Key

OS Location/Qualifiers

OS 4..666

OS /\*tag= a

OS /product= "Ral2-mammaglobin fusion protein"

OS CDS

OS FT

OS FT

OS FT

XX XX

PN WO200125401-A2.  
 XX  
 XX PD 12-APR-2001.  
 XX PF 06-OCT-2000; 2000WO-US027652.  
 XX PR 07-OCT-1999; 99US-0158585P.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Skeiky Y, Guderian J;  
 XX DR WPI; 2001-266299/27.  
 XX DR P-PSDB; AAO22141.  
 XX  
 XX Recombinant nucleic acid molecule for producing high yield expression of  
 XX desired fusion polypeptides, encodes fusion polypeptide comprising  
 XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide.  
 XX  
 XX Disclosure; Fig 5; 39pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2-mammaglobin fusion protein. (Updated  
 CC on 06-AUG-2003 to correct OS field.)  
 XX

SQ Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 4; Length 672;

Best Local Similarity 100.0%; Pred. No. 5.4e-84;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGGCTCCGATTAATTCAGCTGCCAGGTCGCGAGGATTCGCCATTCGATC 60  
 Db 25 ACGGCGGCTCCGATTAATTCAGCTGCCAGGTCGCGAGGATTCGCCATTCGATC 84  
 QY 61 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGGTGGGGTCAACCCCGTTTCA 120  
 Db 85 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGGTGGGGTCAACCCCGTTTCA 144  
 QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGTCTCGACACACGCGACGCGACGA 180  
 Db 145 ATCGGCGCTACCGCTTCTCGGCTTGGTCTCGACACACGCGACGCGACGA 204  
 QY 181 GTCCAAACGCGTGTGGAGGCGTCCCGCGCAAGTCTCGGCACTCTCCACCGCGAGTG 240  
 Db 205 GTCCAAACGCGTGTGGAGGCGTCCCGCGCAAGTCTCGGCACTCTCCACCGCGAGTG 264  
 QY 241 ATCACCGGCTCGACGCGCTCCGATCACTCGGCCACCGCGATGGCGGCGCTTAAC 300  
 Db 265 ATCACCGGCTCGACGCGCTCCGATCACTCGGCCACCGCGATGGCGGCGCTTAAC 324  
 QY 301 GGGCATATCCCGGTAGCTATCTCGTGACCTGGCAACCAAGTCGGCGGCGCGCT 360  
 Db 325 GGGCATATCCCGGTAGCTATCTCGTGACCTGGCAACCAAGTCGGCGGCGCGCT 384  
 QY 361 ACAGGGAACGTCACATTGGCGAGGACCCCGGCC 396  
 Db 395 ACAGGGAACGTCACATTGGCGAGGACCCCGGCC 420

```

RESULT 3
AAZ20206
ID AA220206 standard; DNA; 702 BP.
XX
AC
AAZ20206;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.
XX
KW Tuberculosis; antigen; fusion protein; Mtb24; Ral2; DPPD; diagnosis;
XX therapy; vaccine; immunogen; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9951748-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
XX
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Alderson M, Campos-Neto A;
XX
PI WPI; 1999-601610/51.
XX
DR P-PSDB; AAY32071.
XX
DR
XX
PT New fusion proteins useful for diagnosis, prevention and treatment of
XX tuberculosis.
XX
PS Example; Fig 13A-B; 83pp; English.
XX
CC This DNA sequence includes a coding region for a recombinant
CC Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),
CC termed Mtb24, composed of the antigens Ral2 and DPPD. The DNA is useful
CC for the recombinant production of the fusion protein. Coding sequences
CC for the antigens were modified by PCR in order to facilitate their fusion
CC and subsequent expression of the fusion protein, and then ligated. The
CC invention provides fusion proteins (see AAY32059-71) containing at least
CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC encoding them are useful as vaccines for preventing tuberculosis
CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC progression, and treatment of tuberculosis. They are more effective
CC immunogens than mixtures of the individual protein components
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;
Query Match 100.0%; Score 396; DB 2; Length 702;
Best Local Similarity 100.0%; Pred. No. 5.4e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCGCGTCCGATACCTCCAGCTGCCAGGTCGGGAGGATTCGCCATTCGGATC 60
DB 25 ACGGCGCGTCCGATACCTCCAGCTGCCAGGTCGGGAGGATTCGCCATTCGGATC 84
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTACCCACCGTTTCA 120
DB 85 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTGGGGGTACCCACCGTTTCA 144
QY 121 ATCGGCGCTACCGCTTCTCGGCTGGGTGTCGACACACGACACGCGGCGGCGCAGA 180
DB 145 ATCGGCGCTACCGCTTCTCGGCTGGGTGTCGACACACGCGGCGGCGCAGA 204
QY 181 GTCCACACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
DB 205 GTCCACACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264
QY 241 ATCACCGCGGTGCGACGGGCTCCGATCAACTCGGCCACCGGATCGGCGACGCGCTTAAC 300
DB 265 ATCACCGCGGTGCGACGGGCTCCGATCAACTCGGCCACCGGATCGGCGACGCGCTTAAC 324
QY 301 GGGCATCATCCCGGTGAGGTGATCTCGGTGAGCTCGGCAAAACCAAGTCGGGCGGCGACGGGT 360
DB 325 GGGCATCATCCCGGTGAGGTGATCTCGGTGAGCTCGGCAAAACCAAGTCGGGCGGCGACGGGT 384
QY 361 ACAGGGAACGTGACATTTGGCCGAGGAGCCCGCGGCC 396
DB 385 ACAGGGAACGTGACATTTGGCCGAGGAGCCCGCGGCC 420
RESULT 4
AAL40770
ID AAL40770 standard; DNA; 702 BP.
XX
AC
AAL40770;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ral2-DPPD fusion protein.
XX
KW Ral2; serine protease antigen; Mtb32A; Mycobacterium tuberculosis;
XX vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
XX immunogen; cytokines; gene; ds.
XX
OS Unidentified.
XX
OS Chimeric.
XX
PH Location/Qualifiers
XX FT 4..896
XX FT /*tag= a
XX FT /product= "Ral2-DPPD fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
XX
DR P-PSDB; AAO22139.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
XX desired fusion polypeptides, encodes fusion polypeptide comprising
XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Example 1; Fig 3; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
XX fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
XX kDa C-terminal fragment of serine protease antigen Mtb32A of
XX Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
XX The recombinant fusion nucleic acids and polypeptides are useful for
XX providing stable and high yield expression of fusion polypeptides of both
XX eukaryotic and prokaryotic origin and to encode a protein product for use
XX as an antigen for detecting serum antibodies. The presence of serum
XX antibodies to M. tuberculosis antigens in an individual indicates that
XX the individual is infected with it. The fusion polypeptides are useful as
XX sources of proteins for monitoring binding of serum antibodies to fusion
XX proteins and as an immunogen to induce and/or enhance immune responses.
XX The coding sequences can be ligated with a coding sequence of another.
XX CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
XX represents the DNA encoding the Ral2-DPPD fusion protein
XX
```

```
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;
  Query Match      100.0%; Score 396; DB 4; Length 702;
  Best Local Similarity 100.0%; Pred. No. 5.4e-84;
  Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCCGCTCGATTAATTCAGCTGCCAGGCTGGCAGGATTCGCATTCGGATC 60
Db 25 ACGGCCGCTCGATTAATTCAGCTGCCAGGCTGGCAGGATTCGCATTCGGATC 84
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCGATCGGCTGGGGGTCAACCCACGTTTCAT 120
Db 85 GGGCAGGCGATGGCGATCGCGGCCAGATCGATCGGCTGGGGGTCAACCCACGTTTCAT 144
QY 121 ATCGGGCTACCGCTTCTTCTCGCTTGGGTGTCGACAAACAGCGCAACGGCGCACGA 180
Db 145 ATCGGGCTACCGCTTCTTCTCGCTTGGGTGTCGACAAACAGCGCAACGGCGCACGA 204
QY 181 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGCGAGTG 240
Db 205 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGCGAGTG 264
QY 241 ATACCGGGGTGACGGGCTCCGATCAACTCGGCCACCGCATGGCGAGCGCGCTTAAC 300
Db 265 ATACCGGGGTGACGGGCTCCGATCAACTCGGCCACCGCATGGCGAGCGCGCTTAAC 324
QY 301 GGGCATCATCCGGTGACGTCTCTCGGTGACCTGCGAAACCAAGTCCGGCGGCGACGGGT 360
Db 325 GGGCATCATCCGGTGACGTCTCTCGGTGACCTGCGAAACCAAGTCCGGCGGCGACGGGT 384
QY 361 ACAGGGAAGTGACATTTGGCCGAGGAGCCCGCGGCC 396
Db 385 ACAGGGAAGTGACATTTGGCCGAGGAGCCCGCGGCC 420
RESULT 5
ID ABK14140 standard; DNA; 702 BP.
XX
AC ABK14140;
XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX
DE DNA encoding antigenic fusion protein Ral2-DPPD (Mtb24).
XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW tuberculostatic; immunogen; vaccine; Ral2-DPPD; Mtb24.
XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX
FH Key
FH CDS
FT 1. .702
FT Location/Qualifiers
FT /*tag= a
FT /product= "Mtb24 #1"
FT /partial=
FT /transl_except= (pos:594. .596, aa:Xaa)
FT /notes= "No start or stop codon. Xaa= In frame stop codon"
CDS
FT 2. .702
FT /*tag= b
FT /product= "Mtb24 #2"
FT /partial=
FT /notes= "No start or stop codon"
FT /*transl_except= (pos:263. .265, aa:Xaa) /transl_except=
FT (pos:353. .355, aa:Xaa) /transl_except= (pos:395. .397,
FT aa:Xaa) /transl_except= (pos:470. .472, aa:Xaa)
FT /transl_except= (pos:701. .702, aa:Ser)
FT /note= "This codon has an apparent 1 nucleotide deletion
FT which alters the reading frame. Xaa= In frame stop codon"
FT 3. .701
FT /*tag= c
FT /product= "Mtb24 #3"
FT
```

```
FT
FT /partial
FT /note= "No start or stop codon"
FT /*transl_except= (pos:1. .2, aa:Pro) /transl_except=
FT (pos:39. .41, aa:Xaa) /transl_except= (pos:321. .323,
FT aa:Xaa) /transl_except= (pos:339. .341, aa:Xaa)
FT /*transl_except= (pos:450. .452, aa:Xaa) /transl_except=
FT (pos:621. .623, aa:Xaa)
FT /note= "No start or stop codon. Xaa= In frame stop codon"
XX
PN US2002009459-A1.
XX
PD 24-JAN-2002.
XX
DF 07-APR-1999; 99US-00287849.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (REED/) REED S G.
PA (SKEI/) SKEIKY Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.
XX
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR WPI; 2002-171134/22.
DR P-PSDB; AAU74600, AAU76541, AAU76542.
XX
PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.
XX
PS Example; Fig 13; 62pp; English.
XX
CC The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. This polynucleotide encodes 3 different
CC proteins, each in a different reading frame. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;
  Query Match      100.0%; Score 396; DB 6; Length 702;
  Best Local Similarity 100.0%; Pred. No. 5.4e-84;
  Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCCGCTCGATTAATTCAGCTGCCAGGCTGGCAGGATTCGCATTCGGATC 60
Db 25 ACGGCCGCTCGATTAATTCAGCTGCCAGGCTGGCAGGATTCGCATTCGGATC 84
QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCGATCGGCTGGGGGTCAACCCACGTTTCAT 120
Db 85 GGGCAGGCGATGGCGATCGCGGCCAGATCGATCGGCTGGGGGTCAACCCACGTTTCAT 144
QY 121 ATCGGGCTACCGCTTCTTCTCGCTTGGGTGTCGACAAACAGCGCAACGGCGCACGA 180
Db 145 ATCGGGCTACCGCTTCTTCTCGCTTGGGTGTCGACAAACAGCGCAACGGCGCACGA 204
QY 181 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGCGAGTG 240
Db 205 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGCGAGTG 264
```

QY 241 ATACCGCGGTGACGGCGCTCCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAAC 300  
Db 265 ATACCGCGGTGACGGCGCTCCGATCAACTCGGCCACCGCGATGCGGACGCGCTTAAC 324  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGCAACCAACCAAGTCGGCGGCGCAGCGT 360  
Db 325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGCAACCAACCAAGTCGGCGGCGCAGCGT 384  
QY 361 ACAGGGAACGTGACATTTGGCGCGAGGACCCCGGCC 396  
Db 385 ACAGGGAACGTGACATTTGGCGCGAGGACCCCGGCC 420

## RESULT 6

AAD47078

ID AAD47078 standard; DNA; 1002 BP.

XX AC

XX AAD47078;

XX DT

XX 27-JAN-2003 (first entry)

XX DE

XX Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.

XX KW

XX Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;

XX KW

XX gene; antigen; ds.

XX OS

XX Mycobacterium tuberculosis.

XX OS

XX Synthetic.

XX XX

XX Key

XX Location/Qualifiers

XX CDS

XX 4. .996

XX /\*tag= a

XX /product= "Ra35FLMutSA mutant antigenic protein"

XX XX

XX WO200272792-A2.

XX PN

XX 19-SEP-2002.

XX XX

XX 13-MAR-2002; 2002WO-US008223.

XX XX

XX 13-MAR-2001; 2001US-0275837P.

XX PR

XX (CORI-) CORIXA CORP.

XX XX

XX Skeiky Y, Brannon M, Guderian J;

XX PI

XX WPI; 2002-759844/82.

XX DR

XX P-PSDB; AAE29703.

XX XX

XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,

XX PT

XX M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity

XX PT

XX against pathogenic microorganisms e.g. Leishmania and Mycobacterium

XX PT

XX tuberculosis.

XX XX

XX Disclosure; Page 80-81; 155pp; English.

XX PS

XX The invention relates to a recombinant nucleic acid molecule encoding a

XX CC

XX fusion polypeptide. The recombinant nucleic acid comprises a heterologous

XX CC

XX polynucleotide sequence encoding an antigen or an antigenic fragment from

XX CC

XX polynucleotide sequence encoding an antigenic fragment from

XX CC

XX Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a

XX CC

XX polypeptide or its fragment. The Leishmania polynucleotide is selected

XX CC

XX from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention

XX CC

XX are used in methods for eliciting immune response in mammals. They are

XX CC

XX useful as vaccines to elicit protective immunity against pathogenic

XX CC

XX microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion

Best Local Similarity 100.0%; Pred. No. 5.6e-84; Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGCGCGGTCCGATCACTTCCAGCTGTCCAGGGTGGCAGGATTGCGCATTCGATC 60  
Db 598 ACAGCGCGGTCCGATCACTTCCAGCTGTCCAGGGTGGCAGGATTGCGCATTCGATC 657  
QY 61 GGGCAGGGGATGGCGATCGCGGGCAGATCCGATCGGGTGGGGGTCAACCCACGCTTCA 120  
Db 658 GGGCAGGGGATGGCGATCGCGGGCAGATCCGATCGGGTGGGGGTCAACCCACGCTTCA 717  
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTGACAAACAAACGCGCAACGCG 180  
Db 718 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTGACAAACAAACGCGCAACGCG 777  
QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGGAGCTG 240  
Db 778 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGATCTCCACCGGAGCTG 837  
QY 241 ATCACCGCGGTGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAA 300  
Db 838 ATCACCGCGGTGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAA 897  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGCGT 360  
Db 898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGGGGCGACGCGT 957  
QY 361 ACAGGGAACGTGACATTTGGCGCGAGGACCCCGGCC 396  
Db 958 ACAGGGAACGTGACATTTGGCGCGAGGACCCCGGCC 993

## RESULT 7

AAD47077

ID AAD47077 standard; DNA; 1002 BP.

XX AC

XX AAD47077;

XX XX

XX 27-JAN-2003 (first entry)

XX DT

XX Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.

XX DE

XX Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen; gene;

XX KW

XX ds.

XX XX

XX Mycobacterium tuberculosis.

XX OS

XX Location/Qualifiers

XX PH

XX Key

XX CDS

XX 4. .996

XX /\*tag= a

XX /product= "Ra35 mature antigenic protein"

XX FT

XX /transl\_except= (pos:547. .549, aa:Asp)

XX FT

XX /transl\_except= (pos:550. .552, aa:Ser)

XX XX

XX WO200272792-A2.

XX PN

XX 19-SEP-2002.

XX PD

XX 13-MAR-2002; 2002WO-US008223.

XX XX

XX 13-MAR-2001; 2001US-0275837P.

XX PR

XX (CORI-) CORIXA CORP.

XX XX

XX Skeiky Y, Brannon M, Guderian J;

XX PI

XX WPI; 2002-759844/82.

XX XX

XX P-PSDB; AAE29702.

XX DR

XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,

XX PT

XX M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity

XX PT

XX against pathogenic microorganisms e.g. Leishmania and Mycobacterium

XX PT

XX tuberculosis.

Query Match 100.0%; Score 396; DB 6; Length 1002;

```
XX PS Disclosure; Page 79; 155pp; English.
XX
XX CC The invention relates to a recombinant nucleic acid molecule encoding a
XX CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
XX CC polynucleotide sequence encoding an antigen or an antigenic fragment from
XX CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
XX CC polypeptide or its fragment. The Leishmania polynucleotide is selected
XX CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
XX CC are used in methods for eliciting immune response in mammals. They are
XX CC useful as vaccines to elicit protective immunity against pathogenic
XX CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
XX CC polypeptides are used for enhancing the expression of polynucleotides, as
XX CC in vivo diagnostic agents and for raising antibodies in a non-human
XX CC animal. The invention is used in gene therapy. The present sequence is M.
XX CC tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen encoding
XX CC DNA
XX
XX SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 396; DB 6; Length 1002;
XX Best Local Similarity 100.0%; Pred. No. 5,6e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ACGGCCGCTCCGATTAATCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 60
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 598 ACGGCCGCTCCGATTAATCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 657
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGCTGGGGGTCAACCCCGTTTCA 120
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 658 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGCTGGGGGTCAACCCCGTTTCA 717
XX QY 121 ATCGGGCTACCGCTTCTCTCGGCTTGGGTTGTGACAAACCGGCAACCGGCGACGA 180
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 718 ATCGGGCTACCGCTTCTCTCGGCTTGGGTTGTGACAAACCGGCAACCGGCGACGA 777
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 181 GTCCAAACGCGTGGCGAGCGTCCCGGCGGCAAGTTCGGCATCTCCACCGGCGACGTG 240
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 778 GTCCAAACGCGTGGCGAGCGTCCCGGCGGCAAGTTCGGCATCTCCACCGGCGACGTG 837
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 241 ATCCCGCGTGGCGAGCGTCCGATCAACTCGGCCACCGGCGGCGGCGGCGGCTTAC 300
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 838 ATCCCGCGTGGCGAGCGTCCGATCAACTCGGCCACCGGCGGCGGCGGCGGCTTAC 897
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 301 GGGCATCATCCCGTGGAGTTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGT 360
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 898 GGGCATCATCCCGTGGAGTTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGT 957
XX QY 361 ACAGGGAACGTGACATTGGCGAGGCGACCCCGGCC 396
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 958 ACAGGGAACGTGACATTGGCGAGGCGACCCCGGCC 993
XX
XX RESULT 8
XX AAD28336
XX ID AAD28336 standard; cDNA; 1002 BP.
XX XX
XX AC AAD28336;
XX XX
XX DT 22-APR-2002 (first entry)
XX
XX DE Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX
XX KW Fusion protein; antigen; serological sensitivity; immune response;
XX KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX
XX OS Mycobacterium sp.
XX
XX FH Location/Qualifiers
XX FT 4..996
XX FT /*tag= a
XX FT /product= "Ra35 mature protein"
XX FT /transl_except= (pos:547..549, aa:Asp)
XX FT
```

```
FT XX /transl_except= (pos:550..552, aa:Ser)
FN MO200198460-A2.
XX 27-DEC-2001.
XX 20-JUN-2001; 2001WO-US019959.
XX 20-JUN-2000; 2000US-00597796.
XX 01-FEB-2001; 2001US-0265737P.
XX (CORI-) CORIXA CORP.
XX Skeiky Y, Reed S, Alderson M;
XX WPI: 2002-147798/19.
XX P-PSDB; AAE17566.
XX
XX Composition comprising MTB39 antigen and MTB32A antigen from
XX Mycobacterium species, useful for eliciting immune response in a subject.
XX Disclosure; Page 95; 136pp; English.
XX
XX CC The present invention relates to fusion proteins containing at least two
XX CC Mycobacterium species antigens, nucleotides encoding them and
XX CC compositions comprising such fusion proteins. The present invention
XX CC particularly relates to nucleic acids encoding fusion proteins that
XX CC include two or more individual M. tuberculosis antigens which increase
XX CC the serological sensitivity of sera from individuals infected with
XX CC tuberculosis and methods for their use in diagnosis, prevention and
XX CC treatment of tuberculosis infection. Sequences of the invention are
XX CC useful for eliciting an immune response in a mammal, e.g., human,
XX CC immunised with BCG. They are useful in the diagnosis, treatment and
XX CC prevention of Mycobacterium infection. The fusion proteins and the
XX CC polynucleotides are useful as diagnostic tools in patients infected with
XX CC Mycobacterium, in vitro and in vivo assays for detecting humoral
XX CC antibodies or cell-mediated immunity against M. tuberculosis, for the
XX CC diagnosis of an infection or monitoring of disease progression, as
XX CC immunogens to generate or elicit a protective immune response in a
XX CC patient and for raising anti-M. tuberculosis antibodies in a non-human
XX CC animal. Sequences of the invention are also used as vaccines. MTB32A
XX CC fusion proteins of the invention are useful as in vivo diagnostic agents
XX CC for intradermal skin test. The present sequence is a cDNA encoding
XX CC Mycobacterium species MTB32A (Ra32FL) mature protein
XX
XX SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 396; DB 6; Length 1002;
XX Best Local Similarity 100.0%; Pred. No. 5,6e-84;
XX Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ACGGCCGCTCCGATTAATCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 60
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 598 ACGGCCGCTCCGATTAATCCAGCTGTCACAGGTGGGCGAGGATTCGCCATTCGGATC 657
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGCTGGGGGTCAACCCCGTTTCA 120
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 658 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGCTGGGGGTCAACCCCGTTTCA 717
XX QY 121 ATCGGGCTACCGCTTCTCTCGGCTTGGGTTGTGACAAACCGGCAACCGGCGACGA 180
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 718 ATCGGGCTACCGCTTCTCTCGGCTTGGGTTGTGACAAACCGGCAACCGGCGACGA 777
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 181 GTCCAAACGCGTGGCGAGCGTCCCGGCGGCAAGTTCGGCATCTCCACCGGCGACGTG 240
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 778 GTCCAAACGCGTGGCGAGCGTCCCGGCGGCAAGTTCGGCATCTCCACCGGCGACGTG 837
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 241 ATCCCGCGTGGCGAGCGTCCGATCAACTCGGCCACCGGCGGCGGCGGCGGCTTAC 300
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 838 ATCCCGCGTGGCGAGCGTCCGATCAACTCGGCCACCGGCGGCGGCGGCGGCTTAC 897
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX QY 301 GGGCATCATCCCGTGGAGTTCATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGT 360
XX DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 898 GGGCATCATCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTGGCGGCGACGCGT 957  
AAD28337  
ID AAD28337 standard; CDNA; 1002 BP.  
XX  
AC AAD28337;  
XX  
DT 22-APR-2002 (first entry)  
XX  
XX Mycobacterium species Ra35FLMutSA mutant cDNA.  
DE  
XX Fusion protein; antigen; serological sensitivity; immune response;  
KW tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutain; ss.  
XX  
XX Mycobacterium sp.  
XX  
PH Key Location/Qualifiers  
FT CDS 4..996  
FT /\*tag= a  
FT /product= "Ra35FLMutSA protein"  
XX  
XX WO200198460-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 20-JUN-2001; 2001WO-US019959.  
XX  
XX 20-JUN-2000; 2000US-00597796.  
PR 01-FEB-2001; 2001US-0265737P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skeiky Y, Reed S, Alderson M;  
XX WPI; 2002-147798/19.  
DR P-PSDB; AAE17567.  
XX  
XX Composition comprising MTB39 antigen and MTB32A antigen from  
PT Mycobacterium species, useful for eliciting immune response in a subject.  
XX  
XX Claim 67; Page 96-97; 136pp; English.  
XX  
XX The present invention relates to fusion proteins containing at least two  
CC Mycobacterium species antigens, nucleotides encoding them and  
CC compositions comprising such fusion proteins. The present invention  
CC particularly relates to nucleic acids encoding fusion proteins that  
CC include two or more individual M. tuberculosis antigens which increase  
CC the serological sensitivity of sera from individuals infected with  
CC tuberculosis and methods for their use in diagnosis, prevention and  
CC treatment of tuberculosis infection. Sequences of the invention are  
CC useful for eliciting an immune response in a mammal, e.g., human,  
CC immunised with BCG. They are useful in the diagnosis, treatment and  
CC polynucleotides are useful as diagnostic tools in patients infected with  
CC Mycobacterium, in vitro and in vivo assays for detecting humoral  
CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
CC diagnosis of an infection or monitoring of disease progression, as  
CC immunogens to generate or elicit a protective immune response in a  
CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
CC animal. Sequences of the invention are also used as vaccines. MTB32A  
CC fusion proteins of the invention are useful as in vivo diagnostic agents  
CC for intradermal skin test. The present sequence is a cDNA encoding  
CC Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA  
XX  
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;  
Query Match 100.0%; Score 396; DB 6; Length 1002;

Best Local Similarity 100.0%; Pred. No. 5.6e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCCAGAGGTGGGAGGATTCGCCATCCGATC 60  
DB 598 ACGGCGCGTCCGATAAATTCAGCTGTCCAGAGGTGGGAGGATTCGCCATCCGATC 657  
QY 61 GGGCAGGCGATCGGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTCAT 120  
DB 658 GGGCAGGCGATCGGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTCAT 717  
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACACGCAACGCGCACGA 180  
DB 718 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACGCGCACGA 777  
QY 181 GTCGAACGCGTGGTGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGACGCG 240  
DB 778 GTCGAACGCGTGGTGGGAGCGCTCGGCGGCAAGTCTCGGATCTCCACCGCGACGCG 837  
QY 241 ATCAGCGCGGTGACGCGCGCTCGGATCAACTCGGCCACCGCGATGGCGGACGCGTAAAC 300  
DB 838 ATCAGCGCGGTGACGCGCGCTCGGATCAACTCGGCCACCGCGATGGCGGACGCGTAAAC 897  
QY 301 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCCGGCGCGACGCGT 360  
DB 898 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCCGGCGCGACGCGT 957  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
DB 958 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 993  
RESULT 10  
AA34251  
ID AA34251 standard; DNA; 1068 BP.  
XX  
AC AA34251;  
XX  
DT 06-JUL-1999 (first entry)  
XX  
XX Mycobacterium species nucleic acid sequence 50D.  
DE  
XX Secreted protein; Mycobacterium; primer; PCR; amplification; probe;  
KW hybridisation; detection; vaccine; immunisation; infection; ss.  
XX  
XX Mycobacterium sp.  
XX OS  
XX WO9909186-A2.  
XX  
XX 25-FEB-1999.  
XX  
XX 14-AUG-1998; 98WO-FR001813.  
XX  
XX 14-AUG-1997; 97FR-00010404.  
PR 11-SEP-1997; 97FR-00011325.  
XX  
XX (INSP ) INST PASTEUR.  
XX  
XX Gicquel B, Portnoie D, Lim E, Pelicic V, Guigueno A;  
PI Goguet De La Salmoniere Y;  
XX  
XX WPI; 1999-181045/15.  
DR P-PSDB; AAY05000.  
XX  
XX Mycobacterial DNA vectors containing reporter constructs - for  
PT identifying coding or promoter sequences involved in infection-associated  
PT protein expression.  
XX  
XX Claim 22; Fig 50D; 309pp; French.  
XX  
CC Sequences AA34001-X34252 represent nucleic acids encoding secreted  
CC proteins from various Mycobacterium species microorganisms. The  
CC nucleotide sequences can be used as primers and probes for methods for

CC detecting and identifying mycobacteria, especially belonging to the M.  
 CC tuberculosis complex. The encoded proteins can be used in vaccines for  
 CC immunisation against a bacterial or viral infection  
 XX  
 SQ Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 396; DB 2; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-84;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGGCGCGTCCGATTAACCTTCAGCTGTCACAGGGTGGGAGGATTCCGCAATCCGATC 60  
 Db 670 ACGGCGCGTCCGATTAACCTTCAGCTGTCACAGGGTGGGAGGATTCCGCAATCCGATC 729  
 QY 61 GGGCAGGCGATCGGATCGGGGCGAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 120  
 Db 730 GGGCAGGCGATCGGATCGGGGCGAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 789  
 QY 121 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTTGCGACAAACGCGCAACGGGCGCAGA 180  
 Db 790 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTTGCGACAAACGCGCAACGGGCGCAGA 849  
 QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCGAAGTCTCGGATCTCCACGGGCGACGTG 240  
 Db 850 GTCCAAACGGTGGTGGGAGCGCTCGGGGCGAAGTCTCGGATCTCCACGGGCGACGTG 909  
 QY 241 ATCACCGCGGTTCGACGGGCGCTCCGATCAACTCGGCGACCGGATGGCGGACGGCTTAAC 300  
 Db 910 ATCACCGCGGTTCGACGGGCGCTCCGATCAACTCGGCGACCGGATGGCGGACGGCTTAAC 969  
 QY 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGGCAACCAAGTCGGGCGGACGGCT 360  
 Db 970 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGGCAACCAAGTCGGGCGGACGGCT 1029  
 QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
 Db 1030 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 1065  
 RESULT 11  
 AAX34252  
 ID AAX34252 standard; DNA; 1143 BP.  
 XX  
 AC AAX34252;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE Mycobacterium species nucleic acid sequence 50F.  
 XX  
 KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;  
 KW hybridisation; detection; vaccine; immunisation; infection; ss.  
 XX  
 OS Mycobacterium sp.  
 XX  
 PN WO9909186-A2.  
 XX  
 PD 25-FEB-1999.  
 XX  
 PF 14-AUG-1998; 98WO-FR001813.  
 XX  
 PR 14-AUG-1997; 97FR-00010404.  
 PR 11-SEP-1997; 97FR-00011325.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Gicquel B, Portnoie D, Lim E, Pellicic V, Guigneno A;  
 PI Gouget De La Salmoniere Y;  
 XX  
 DR WPI; 1999-181045/15.  
 DR P-PSDB; AAY04830.  
 XX  
 PT Mycobacterial DNA vectors containing reporter constructs - for  
 PT identifying coding or promoter sequences involved in infection-associated

PT protein expression.  
 XX Claim 22; Fig 50F; 309pp; French.  
 XX  
 CC Sequences AAX34001-X34252 represent nucleic acids encoding secreted  
 CC proteins from various Mycobacterium species microorganisms. The  
 CC nucleotide sequences can be used as primers and probes for methods for  
 CC detecting and identifying mycobacteria, especially belonging to the M.  
 CC tuberculosis complex. The encoded proteins can be used in vaccines for  
 CC immunisation against a bacterial or viral infection  
 XX  
 SQ Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 396; DB 2; Length 1143;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-84;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGGCGCGTCCGATTAACCTTCAGCTGTCACAGGGTGGGAGGATTCCGCAATCCGATC 60  
 Db 745 ACGGCGCGTCCGATTAACCTTCAGCTGTCACAGGGTGGGAGGATTCCGCAATCCGATC 804  
 QY 61 GGGCAGGCGATCGGATCGGGGCGAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 120  
 Db 805 GGGCAGGCGATCGGATCGGGGCGAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 864  
 QY 121 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTTGCGACAAACGCGCAACGGGCGCAGA 180  
 Db 865 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTTGCGACAAACGCGCAACGGGCGCAGA 924  
 QY 181 GTCCAAACGGTGGTGGGAGCGCTCGGGGCGAAGTCTCGGATCTCCACGGGCGACGTG 240  
 Db 925 GTCCAAACGGTGGTGGGAGCGCTCGGGGCGAAGTCTCGGATCTCCACGGGCGACGTG 984  
 QY 241 ATCACCGCGGTTCGACGGGCGCTCCGATCAACTCGGCGACCGGATGGCGGACGGCTTAAC 300  
 Db 985 ATCACCGCGGTTCGACGGGCGCTCCGATCAACTCGGCGACCGGATGGCGGACGGCTTAAC 1044  
 QY 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGGCAACCAAGTCGGGCGGACGGCT 360  
 Db 1045 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGGCAACCAAGTCGGGCGGACGGCT 1104  
 QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
 Db 1105 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 1140  
 RESULT 12  
 AAL40771  
 ID AAL40771 standard; DNA; 1742 BP.  
 XX  
 AC AAL40771;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Nucleotide sequence encoding Ral2-WT1 fusion protein.  
 XX  
 KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
 KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
 KW immunogen; cytokine; gene; ds.  
 XX  
 OS Unidentified.  
 OS Chimeric.  
 XX  
 PH Key Location/Qualifiers  
 PH CDS 4..1740  
 ET /tag= a  
 ET /product= "Ral2-WT1 fusion protein"  
 XX  
 XX WO200125401-A2.  
 XX  
 PN 12-APR-2001.  
 XX  
 PD 06-OCT-2000; 2000WO-US027652.  
 PF



XX 07-OCT-1999; 99US-0158585P.  
XX (CORI-) CORIXA CORP.  
XX Skeiky Y, Guderian J;  
XX WPI; 2001-266299/27.  
XX P-PSDB; AAO22140.  
XX  
XX Recombinant nucleic acid molecule for producing high yield expression of  
XX desired fusion polypeptides, encodes fusion polypeptide comprising  
XX Mycobacterium tuberculosis coding sequence and heterologous polypeptide.  
XX  
XX Disclosure; Fig 4; 39pp; English.  
XX  
XX The invention relates to a recombinant nucleic acid molecule encoding a  
XX fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
XX kDa C-terminal fragment of serine protease antigen MTB32A of  
XX Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
XX The recombinant fusion nucleic acids and polypeptides are useful for  
XX providing stable and high yield expression of fusion polypeptides of both  
XX eukaryotic and prokaryotic origin and to encode a protein product for use  
XX as an antigen for detecting serum antibodies. The presence of serum  
XX antibodies to M. tuberculosis antigens in an individual indicates that  
XX the individual is infected with it. The fusion polypeptides are useful as  
XX sources of proteins for monitoring binding of serum antibodies to fusion  
XX proteins and as an immunogen to induce and/or enhance immune responses.  
XX The coding sequences can be ligated with a coding sequence of another  
XX molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
XX can be used in vivo as a DNA vaccine. This polynucleotide sequence  
XX represents the DNA encoding the Ral2-WT1 fusion protein  
XX  
XX Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 396; DB 4; Length 1742;  
XX Best Local Similarity 100.0%; Pred. No. 5.7e-84; Indels 0; Gaps 0;  
XX Matches 396; Conservative 0; Mismatches 0;  
XX  
XX 1 ACGGCGCGTCCGATAAATTCAGCTGCCAGAGGTGGCGAGGATTCGCCATTCCGATC 60  
XX 25 ACGGCGCGTCCGATAAATTCAGCTGCCAGAGGTGGCGAGGATTCGCCATTCCGATC 84  
XX 61 GGGCAGGCGATGCGGATCGCGGCCAGATCCGATCGGTTGGGGGTACCCACCGTTTCAT 120  
XX 85 GGGCAGGCGATGCGGATCGCGGCCAGATCCGATCGGTTGGGGGTACCCACCGTTTCAT 144  
XX 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTGACAAACACGCGGACGCGACGA 180  
XX 145 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTGACAAACACGCGGACGCGACGA 204  
XX 181 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGACGCGT 240  
XX 205 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGACGCGT 264  
XX 241 ATCCGCGGTGACGCGCTCCGATCACTCGGCGGACGCGGATCGGCGGACGCGTTAAC 300  
XX 265 ATCCGCGGTGACGCGCTCCGATCACTCGGCGGACGCGGATCGGCGGACGCGTTAAC 324  
XX 301 GGGCATCATCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 360  
XX 325 GGGCATCATCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 384  
XX 361 ACAGGACGTCACATTTGCCAGGACCCCGGCC 396  
XX 385 ACAGGACGTCACATTTGCCAGGACCCCGGCC 420  
XX  
XX RESULT 13  
XX ID ADA26353  
XX standard; DNA; 2181 BP.  
XX  
XX ADA26353;

XX 20-NOV-2003 (first entry)  
XX Mycobacterium MTB32-MTB39F fusion protein encoding DNA.  
XX  
XX ds; gene; fusion polypeptide; MTB32A; MTB39; antigen; MTB32A; MTB39;  
XX MTB85A; tuberculosis; tuberculostatic; gene therapy; vaccine.  
XX  
XX Chimeric.  
XX Mycobacterium sp.  
XX  
XX Key Location/Qualifiers  
XX CDS 4..2175  
XX /\*tag= a  
XX /product= "MTB32-MTB39F fusion polypeptide"  
XX  
XX WO2003070187-A2.  
XX 28-AUG-2003.  
XX 18-FEB-2003; 2003WO-US004903.  
XX 15-FEB-2002; 2002US-0357351P.  
XX (CORI-) CORIXA CORP.  
XX Skeiky Y, Guderian J, Reed S;  
XX WPI; 2003-597554/66.  
XX P-PSDB; ADA26354.  
XX  
XX New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
XX and MTB39 antigens, with or without the MTB85A antigen, from a  
XX Mycobacterium species, useful for diagnosing, preventing and/or treating  
XX tuberculosis infection.  
XX  
XX Claim 1; Fig 1; 112pp; English.  
XX  
XX The invention relates to a novel nucleic acid encoding a fusion  
XX polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
XX MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
XX A polypeptide of the invention has tuberculostatic activity. A  
XX polynucleotide of the invention may have a use in gene therapy, and as a  
XX vaccine. The methods and compositions of the present invention are useful  
XX for diagnosing, preventing and/or treating tuberculosis infection. The  
XX present sequence is used in the exemplification of the invention.  
XX  
XX Sequence 2181 BP; 341 A; 693 C; 792 G; 355 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 396; DB 8; Length 2181;  
XX Best Local Similarity 100.0%; Pred. No. 5.8e-84; Indels 0; Gaps 0;  
XX Matches 396; Conservative 0; Mismatches 0;  
XX  
XX 1 ACGGCGCGTCCGATAAATTCAGCTGCCAGAGGTGGCGAGGATTCGCCATTCCGATC 60  
XX 598 ACGGCGCGTCCGATAAATTCAGCTGCCAGAGGTGGCGAGGATTCGCCATTCCGATC 657  
XX 61 GGGCAGGCGATGCGGATCGCGGCCAGATCCGATCGGTTGGGGGTACCCACCGTTTCAT 120  
XX 658 GGGCAGGCGATGCGGATCGCGGCCAGATCCGATCGGTTGGGGGTACCCACCGTTTCAT 717  
XX 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTGACAAACACGCGGACGCGACGA 180  
XX 718 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTGTGACAAACACGCGGACGCGACGA 777  
XX 181 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGACGCGT 240  
XX 778 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGACGCGT 837  
XX 241 ATCACC CGGTCCAGCGCGCTCCGATCAACTCGGCGGACGCGGATCGGCGGACGCGTTAAC 300  
XX 838 ATCACC CGGTCCAGCGCGCTCCGATCAACTCGGCGGACGCGGATCGGCGGACGCGTTAAC 897



QY 301 GGGCATCATCCGGTGAGTTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGT 360  
 XX  
 Db 898 GGGCATCATCCGGTGAGTTCATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGT 957  
 QY 361 ACAGGAACGTGACATTGGCCGAGGACCCCGGCC 396  
 Db 958 ACAGGAACGTGACATTGGCCGAGGACCCCGGCC 993

RESULT 14

AAD47084  
 ID AAD47084 standard; DNA; 2190 BP.  
 AC  
 XX  
 AC AAD47084;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 27-JAN-2003 (first entry)  
 XX  
 DE Mycobacterium sp. MTB72FmutSA fusion protein encoding DNA.  
 XX  
 XX Vaccine; immunity; diagnostic agent; gene therapy; TBH9; antigen;  
 KW Ra35MutSA; Ra12; MTB72MutSA; chimeric; Gene; ds.  
 XX  
 OS Mycobacterium sp.  
 OS Mycobacterium tuberculosis.  
 OS Chimeric.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2190  
 FT /tag= a  
 FT /product= "MTB72FmutSA fusion protein"

WO200272792-A2.

XX 19-SEP-2002.  
 XX 13-MAR-2002; 2002WO-US008223.  
 XX 13-MAR-2001; 2001US-0275837P.  
 XX (CORI-) CORIXA CORP.  
 XX  
 XX Skeiky Y, Brannon M, Guderian J;  
 XX WPI; 2002-759844/82.  
 DR P-PSDB; AAE29709.  
 XX

XX New recombinant nucleic acid molecule comprising a Leishmania TSA. LeIF,  
 FT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity  
 FT against pathogenic microorganisms e.g. Leishmania and Mycobacterium  
 FT tuberculosis.  
 XX

Disclosure; Page 92-93; 155pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
 CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
 CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
 CC are used in methods for eliciting immune response in mammals. They are  
 CC useful as vaccines to elicit protective immunity against pathogenic  
 CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
 CC polypeptides are used for enhancing the expression of polynucleotides, as  
 CC in vivo diagnostic agents and for raising antibodies in a non-human  
 CC animal. The invention is used in gene therapy. The present sequence is a  
 CC DNA encoding MTB72F fusion protein. This fusion protein comprises  
 CC Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and TBH9  
 CC protein from Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS  
 CC field)

Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 6; Length 2190;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-84;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACAGCGCGCTCCGTAACATTCAGCTGTCCAGGGTGGGCGAGGATTCCGCAATCCGATC 60  
 Db 22 ACAGCGCGCTCCGTAACATTCAGCTGTCCAGGGTGGGCGAGGATTCCGCAATCCGATC 81  
 QY 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTGACCCACCGTTTCAT 120  
 Db 82 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTGACCCACCGTTTCAT 141  
 QY 121 ATCGGGCCCTACCGCTTCTCGCTTGGGTGTGTCGCAACCAACGCGACGCGACGA 180  
 Db 142 ATCGGGCCCTACCGCTTCTCGCTTGGGTGTGTCGCAACCAACGCGACGCGACGA 201  
 QY 181 GTCCAAACGGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
 Db 202 GTCCAAACGGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261  
 QY 241 ATCAACCGCGTGCAGCGCGCTCCGATCAACTCGGCACCGCGATGGCGGCGGCTTAAC 300  
 Db 262 ATCAACCGCGTGCAGCGCGCTCCGATCAACTCGGCACCGCGATGGCGGCGGCTTAAC 321  
 QY 301 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 360  
 Db 322 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 381  
 QY 361 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
 Db 382 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 417

RESULT 15

AAD28343  
 ID AAD28343 standard; DNA; 2190 BP.

AC AAD28343;

XX 22-APR-2002 (first entry)

DE Mycobacterium species MTB72FmutSA fusion protein encoding DNA.

XX Fusion protein; antigen; serological sensitivity; immune response;  
 KW tuberculosis; infection; vaccine; MTB72FmutSA; Ra12-TbH9-Ra35MutSA;  
 KW mutant; mutein; ds.

OS Mycobacterium sp.

XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2190  
 FT /tag= a  
 FT /product= "MTB72FmutSA fusion protein"  
 FT misc\_feature 22..417  
 FT /tag= b  
 FT /note= "Ra12 DNA fragment"  
 FT misc\_feature 424..1596  
 FT /tag= c  
 FT /note= "TBH9FL DNA fragment"  
 FT misc\_feature 1603..2187  
 FT /tag= d  
 FT /note= "Ra35 DNA fragment"  
 FT mutation replace(2128, T)  
 FT /tag= e

WO200198460-A2.

XX 27-DEC-2001.  
 XX 20-JUN-2001; 2001WO-US019959.  
 XX 20-JUN-2000; 2000US-00597796.  
 PR 01-FEB-2001; 2001US-0285737P.

```
XX PA (CORI-) CORIXA CORP.
XX PI
XX PI Skeiky Y, Reed S, Alderson M;
XX DR WPI; 2002-147798/19.
XX DR P-PSDB; AAE17573.
XX PT Composition comprising MTB39 antigen and MTB32A antigen from
XX PT Mycobacterium species, useful for eliciting immune response in a subject.
XX PS Claim 81; Page 108-109; 136pp; English.
XX CC The present invention relates to fusion proteins containing at least two
XX CC Mycobacterium species antigens, nucleotides encoding them and
XX CC compositions comprising such fusion proteins. The present invention
XX CC particularly relates to nucleic acids encoding fusion proteins that
XX CC include two or more individual M. tuberculosis antigens which increase
XX CC the serological sensitivity of sera from individuals infected with
XX CC tuberculosis and methods for their use in diagnosis, prevention and
XX CC treatment of tuberculosis infection. Sequences of the invention are
XX CC useful for eliciting an immune response in a mammal, e.g., human,
XX CC immunised with BCG. They are useful in the diagnosis, treatment and
XX CC prevention of Mycobacterium infection. The fusion proteins and the
XX CC polynucleotides are useful as diagnostic tools in patients infected with
XX CC Mycobacterium, in vitro and in vivo assays for detecting humoral
XX CC antibodies or cell-mediated immunity against M. tuberculosis, for the
XX CC diagnosis of an infection or monitoring of disease progression, as
XX CC immunogens to generate or elicit a protective immune response in a
XX CC patient and for raising anti-M. tuberculosis antibodies in a non-human
XX CC animal. Sequences of the invention are also used as vaccines. MTB32A
XX CC fusion proteins of the invention are useful as in vivo diagnostic agents
XX CC for intradermal skin test. The present sequence is a DNA encoding
XX CC Mycobacterium species MTB72FmutSA (Ral2-TbH9-Ra35MutSA) mutant fusion
XX CC protein
XX SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;
Query Match 100.0%; Score 396; DB 6; Length 2190;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCGCGTCCGATACCTCCAGTGTCCAGGTTGGGAGGATTCGGCATTCGGATC 60
DB 22 ACGGCGCGTCCGATACCTCCAGTGTCCAGGTTGGGAGGATTCGGCATTCGGATC 81
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGGTTTCAT 120
DB 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGGTTTCAT 141
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACACGCGACGCGACGCA 180
DB 142 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACACGCGACGCGACGCA 201
QY 181 GTCCAAACGCGTGTTCGGGAGGCTCCGCGGCAAGTCTTCGGCATCTCCACCGGCGAGTG 240
DB 202 GTCCAAACGCGTGTTCGGGAGGCTCCGCGGCAAGTCTTCGGCATCTCCACCGGCGAGTG 261
QY 241 ATCAACGCGGTTCAGCGGTTCGGATCACTCGGCGCAGGATCGGCGAGCGGCTTAAC 300
DB 262 ATCAACGCGGTTCAGCGGTTCGGATCACTCGGCGCAGGATCGGCGAGCGGCTTAAC 321
QY 301 GGGCATCATCCGGTGAGTCACTTCGGTGACCTCGGCAAAACCAAGTCGGGCGGCGACGCGT 360
DB 322 GGGCATCATCCGGTGAGTCACTTCGGTGACCTCGGCAAAACCAAGTCGGGCGGCGAGT 381
QY 361 ACAGGGAAGTGAACATGGCCGAGGAGCCCGCGCC 396
DB 382 ACAGGGAAGTGAACATGGCCGAGGAGCCCGCGCC 417
RESULT 16
AAL40773
```

```
ID XX AAL40773 standard; DNA; 2191 BP.
AC XX AAL40773;
XX DT 03-OCT-2002 (first entry)
XX DE Nucleotide sequence encoding Ral2-H9-32A fusion protein.
XX KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
XX KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
XX KW immunogen; cytokine; gene; ds.
XX OS Unidentified.
XX OS Chimeric.
XX FH Key Location/Qualifiers
XX FT 1..2190
XX FT /tag=a
XX FT /product="Ral2-H9-32A fusion protein"
XX PN W0200125401-A2.
XX PD 12-APR-2001.
XX PF 06-OCT-2000; 2000WO-US027652.
XX PR 07-OCT-1999; 99US-0158585P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky Y, Guderian J;
XX DR WPI; 2001-266299/27.
XX DR P-PSDB; AAO22142.
XX PT Recombinant nucleic acid molecule for producing high yield expression of
XX PT desired fusion polypeptides, encodes fusion polypeptide comprising
XX PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX PS Disclosure; Fig 6; 39pp; English.
XX CC The invention relates to a recombinant nucleic acid molecule encoding a
XX CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
XX CC kDa C-terminal fragment of serine protease antigen MTB32A of
XX CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
XX CC The recombinant fusion nucleic acids and polypeptides are useful for
XX CC providing stable and high yield expression of fusion polypeptides of both
XX CC eukaryotic and prokaryotic origin and to encode a protein product for use
XX CC as an antigen for detecting serum antibodies. The presence of serum
XX CC antibodies to M. tuberculosis antigens in an individual indicates that
XX CC the individual is infected with it. The fusion polypeptides are useful as
XX CC sources of proteins for monitoring binding of serum antibodies to fusion
XX CC proteins and as an immunogen to induce and/or enhance immune responses.
XX CC The coding sequences can be ligated with a coding sequence of another
XX CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
XX CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
XX CC represents the DNA encoding the Ral2-H9-32A fusion protein
XX SQ Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 1 Other;
Query Match 100.0%; Score 396; DB 4; Length 2191;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCGCGTCCGATACCTCCAGTGTTCAGGTTGGGAGGATTCGGCATTCGGATC 60
DB 22 ACGGCGCGTCCGATACCTCCAGTGTTCAGGTTGGGAGGATTCGGCATTCGGATC 81
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGGTTTCAT 120
DB 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGGTTTCAT 141
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACACGCGACGCGACGCA 180
```

```

Db 142 ATCGGGCTACGGCTTCTCGGTTGGTGTGTGACAAACGCGCAACGGCGGACGA 201
Qy 181 GTCCAAACGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 240
Db 202 GTCCAAACGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 261
Qy 241 ATCACCAGGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 300
Db 262 ATCACCAGGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 321
Qy 301 GGGCATCATCCGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 360
Db 322 GGGCATCATCCGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 381
Qy 361 ACAGGGAACGTGACATTCGCGAGGCGACCCCGCGCC 396
Db 382 ACAGGGAACGTGACATTCGCGAGGCGACCCCGCGCC 417

```

```

RESULT 17
ID ABK14128 standard; DNA; 2286 BP.
XX AC ABK14128;
XX XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX XX
DE DNA encoding antigenic fusion protein Ral2-TbH9-Ra35 (Mtb32-Mtb39).
XX XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW tuberculosatic; immunogen; vaccine; Mtb32-Mtb39; Ral2; TbH9; Ra35;
KW Ral2-TbH9-Ra35.
XX XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX XX
FH Key Location/Qualifiers
CDS 42..2231
FT /product= "Mtb32-Mtb39"
FT /transl_except= (pos:498..506, aa:Asn-Ala)
FT /transl_except= (pos:597..605, aa:Ala-Gln)
FT /transl_except= (pos:798..802, aa:Ala)
FT /note= "This codon has an apparent 2 nucleotide insertion
FT which alters the reading frame"
XX XX
PN US2002009459-A1.
XX XX
XX 24-JAN-2002.
XX XX
XX 07-APR-1999; 99US-00287849.
XX XX
XX 13-MAR-1997; 97US-00818112.
XX 01-OCT-1997; 97US-00942578.
XX 18-FEB-1998; 98US-00025197.
XX 07-APR-1998; 98US-00056556.
XX 30-DEC-1998; 98US-00223040.
XX XX
XX (REED/) REED S G.
XX (SKEI/) SKEIKY Y A.
XX (DILL/) DILLON D C.
XX (ALDE/) ALDERSON M.
XX (CAMP/) CAMPOS-NETO A.
XX XX
XX Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX WFI; 2002-171134/22.
XX P-PSDB; AAU74588.
XX XX
XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for
XX diagnosing, treating or preventing M. tuberculosis infection,
PT

```

```

PT particularly as vaccine for treating or preventing tuberculosis.
XX Example; Fig 1; 62pp; English.
XX The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. (Updated on 29-AUG-2003 to standardise
CC OS field)
XX XX
SQ Sequence 2286 BP; 373 A; 719 C; 807 G; 381 T; 0 U; 6 Other;
Query Match 100.0%; Score 396; DB 6; Length 2286;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACGGCGCGCTCCGATTAATTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 60
Db 63 ACGGCGCGCTCCGATTAATTCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGGATC 122
Qy 61 GGGCAGCGCATGGCGATCGGGCCAGATCCGATCGGTGGGGGTTCACCCACCGTTTCA 120
Db 123 GGGCAGCGCATGGCGATCGGGCCAGATCCGATCGGTGGGGGTTCACCCACCGTTTCA 182
Qy 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCACAAACGCGCAACGGCGACGA 180
Db 183 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTGTCACAAACGCGCAACGGCGACGA 242
Qy 181 GTCCAAACGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 240
Db 243 GTCCAAACGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 302
Qy 241 ATCACCAGGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 300
Db 303 ATCACCAGGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 362
Qy 301 GGGCATCATCCGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 360
Db 363 GGGCATCATCCGGTGTGCGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACGGCGAGTG 422
Qy 361 ACAGGGAACGTGACATTCGCGAGGCGACCCCGCGCC 396
Db 423 ACAGGGAACGTGACATTCGCGAGGCGACCCCGCGCC 458

```

```

RESULT 18
AAZ20194
ID AAZ20194 standard; DNA; 2287 BP.
XX XX
XX AC AAZ20194;
XX XX
DT 17-JAN-2000 (first entry)
XX XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb32A DNA.
XX XX
KW Tuberculosis; antigen; fusion protein; Mtb32A; Ral2; TbH9; Ra35;
KW diagnosis; therapy; vaccine; immunogen; ss.
XX XX
OS Mycobacterium tuberculosis.
XX XX
FH Key Location/Qualifiers
CDS 42..2231
FT /*tag= a
XX XX
XX WO9951748-A2.
XX 14-OCT-1999.
PT

```

```
XX PF 07-APR-1999; 99WO-US007717.
XX PR 07-APR-1998; 98US-00056556.
XX PR 30-DEC-1998; 98US-00223040.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Alderson M, Campos-Neto A;
XX DR WPI; 1999-601610/51.
XX DR P-PSDB; AAY32059.
XX PT New fusion proteins useful for diagnosis, prevention and treatment of
XX FT tuberculosis.
XX PS Example; Fig 1A-B; 83pp; English.
XX CC This DNA sequence includes a coding region for a recombinant
XX CC Mycobacterium tuberculosis tri-antigen fusion protein (see AAY32059),
XX CC termed MtB32A, composed of the antigens Ra12, TbH9 and Ra35. The DNA is
XX CC useful for the recombinant production of the fusion protein. Coding
XX CC sequences for the antigens were modified by PCR in order to facilitate
XX CC their fusion and subsequent expression of the fusion protein. 3 Coding
XX CC sequences for Ra12, TbH9 and Ra25 were ligated to encode MtB32A. The
XX CC invention provides fusion proteins (see AAY32059-71) containing at least
XX CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
XX CC encoding them are useful as vaccines for preventing tuberculosis
XX CC (Claimed), for diagnosis (via in vitro assays or intradermal skin tests
XX CC for detection of anti-M. tuberculosis antibodies), monitoring of disease
XX CC progression, and treatment of tuberculosis. They are more effective
XX CC immunogens than mixtures of the individual protein components
XX SQ Sequence 2287 BP; 374 A; 723 C; 805 G; 382 T; 0 U; 3 Other;

Query Match 100.0%; Score 396; DB 2; Length 2287;
Best Local Similarity 100.0%; Pred. No. 5.8e-84;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGCCAGGGTGGCAGGATTCGCATTCGGATC 60
DB 63 ACGGCGCGTCCGATAAATTCAGCTGCCAGGGTGGCAGGATTCGCATTCGGATC 122
QY 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 120
DB 123 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 182
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACACACGCGACGGCGACGA 180
DB 183 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACACACGCGACGGCGACGA 242
QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGCGCAAGTCTCGCATCTCCACCGGCGAGTG 240
DB 243 GTCCAAACGCGTGGTGGGAGCGCTCCGCGCAAGTCTCGCATCTCCACCGGCGAGTG 302
QY 241 ATCACCGGGTGGGAGCGCTCCGATCACTCGGCCACCGGATGGCGGCGCGCTTAAC 300
DB 303 ATCACCGGGTGGGAGCGCTCCGATCACTCGGCCACCGGATGGCGGCGCGCTTAAC 362
QY 301 GGGCATCATCCGGTGACGTGATCTCGGTGACCTGGCAAAACCAAGTGGGGCGGACGCGT 360
DB 363 GGGCATCATCCGGTGACGTGATCTCGGTGACCTGGCAAAACCAAGTGGGGCGGACGCGT 422
QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396
DB 423 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 458
```

RESULT 19

AAD47083

ID AAD47083 standard; DNA; 2287 BP.

XX

AC AAD47083;

```
XX 29-AUG-2003 (revised)
XX DT 27-JAN-2003 (first entry)
XX DE Mycobacterium sp. MTB72F fusion protein encoding DNA.
XX KW Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen; Ra35;
XX KW Ra12; MTB72F; chimeric; gene; ds.
XX OS Mycobacterium sp.
XX OS Mycobacterium tuberculosis.
XX OS Chimeric.
XX PH Key Location/Qualifiers
XX FT CDS 42..2231
XX FT /*tag= a
XX FT /product= "MTB72F fusion protein"
XX PN WO200272792-A2.
XX PD 19-SEP-2002.
XX PF 13-MAR-2002; 2002WO-US008223.
XX PR 13-MAR-2001; 2001US-0275837P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky Y, Brannon M, Guderian J;
XX DR WPI; 2002-759844/82.
XX DR P-PSDB; AAS29708.
XX PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
XX PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
XX PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
XX PT tuberculosis.
XX PS Disclosure; Page 87-90; 155pp; English.
XX CC The invention relates to a recombinant nucleic acid molecule encoding a
XX CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
XX CC polynucleotide sequence encoding an antigen or an antigenic fragment from
XX CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
XX CC polypeptide or its fragment. The Leishmania polynucleotide is selected
XX CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
XX CC are used in methods for eliciting immune response in mammals. They are
XX CC useful as vaccines to elicit protective immunity against pathogenic
XX CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
XX CC polypeptides are used for enhancing the expression of polynucleotides, as
XX CC in vivo diagnostic agents and for raising antibodies in a non-human
XX CC animal. The invention is used in gene therapy. The present sequence is a
XX CC DNA encoding MTB72F fusion protein. This fusion protein comprises Ra12
XX CC and Ra35 protein from Mycobacterium tuberculosis and TbH9 protein from
XX CC Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS field)
XX SQ Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;
```

Query Match 100.0%; Score 396; DB 6; Length 2287;  
Best Local Similarity 100.0%; Pred. No. 5.8e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ACGGCGCGTCCGATAAATTCAGCTGCCAGGGTGGCAGGATTCGCATTCGGATC 60
DB 63 ACGGCGCGTCCGATAAATTCAGCTGCCAGGGTGGCAGGATTCGCATTCGGATC 122
QY 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 120
DB 123 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 182
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACACACGCGACGGCGACGA 180
DB 183 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACACACGCGACGGCGACGA 242
```





Db	325	GGGCATCATCCGGTGAGTTCATCTCGGTGACCTGGCAACCAAGTCCGGGGGCAACGGCGT	384
Qy	361	ACAGGGAACGTTGACATTCGGCCGAGGAGACCCCGGCGC	396
Db	385	ACAGGGAACGTTGACATTCGGCCGAGGAGACCCCGGCGC	420
RESULT 23			
AD26358			
ID	ADA26358	standard; DNA; 2637 BP.	
XX	ADA26358;		
AC	ADA26358;		
DT	20-NOV-2003	(first entry)	
XX	Mycobacterium MTB72F-Erd14	(fusion MTB89F) protein encoding DNA.	
DE	ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;		
XX	tuberculosis; tuberculostatic; gene therapy; vaccine.		
KW	Chimeric.		
OS	Mycobacterium sp.		
XX	Key	Location/Qualifiers	
PH	CDS	4...2631	
FT	FT	/*tag= a	
FT	FT	/product= "MTB72F-Erd14 (fusion MTB89F)"	
XX	WO2003070187-A2.		
PN	28-AUG-2003.		
XX	18-FEB-2003; 2003WO-US004903.		
XX	15-FEB-2002; 2002US-0357351P.		
XX	(CORI-) CORIXA CORP.		
XX	Skeiky Y, Guderian J, Reed S;		
PI	WPI; 2003-697554/66.		
DR	P-PSDB; ADA26365.		
XX	New isolated nucleic acid encoding a fusion polypeptide with the MTB32A		
PT	and MTB39 antigens, with or without the MTB85A antigen, from a		
PT	Mycobacterium species, useful for diagnosing, preventing and/or treating		
PT	tuberculosis infection.		
XX	Claim 84; Fig 6; 112pp; English.		
PS	The invention relates to a novel nucleic acid encoding a fusion		
CC	polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and		
CC	MTB85A antigen from a Mycobacterium species of the tuberculosis complex.		
CC	A polypeptide of the invention has tuberculostatic activity. A		
CC	polynucleotide of the invention may have a use in gene therapy, and as a		
CC	vaccine. The methods and compositions of the present invention are useful		
CC	for diagnosing, preventing and/or treating tuberculosis infection. The		
CC	present sequence is used in the exemplification of the invention.		
XX	Sequence 2637 BP; 428 A; 840 C; 928 G; 441 T; 0 U; 0 Other;		
SQ	Query Match	100.0%; Score 396; DB 8; Length 2637;	
	Best Local Similarity	100.0%; Pred. No. 5.9e-84;	
	Matches 396; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	ACGGCCGCGTCCGATTAACCTCCAGCTGTCACGGGTGGGCGAGGATTCCGCAATTCGATC	60
Db	25	ACGGCCGCGTCCGATTAACCTCCAGCTGTCACGGGTGGGCGAGGATTCCGCAATTCGATC	84
Qy	61	GGGACGCGATGCGATTCGGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTGAT	120
Db	85	GGGACGCGATGCGATTCGGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTGAT	144

Qy	121	ATCGGGCTTACCGCTTCTCTCGGCTTGGTGTGTCGACAAACGCAACGGCGCACGA	180
Db	145	ATCGGGCTTACCGCTTCTCTCGGCTTGGTGTGTCGACAAACGCAACGGCGCACGA	204
Qy	181	GTCCAAACGCGTGTGTCGGAGCGCTCCGGGGCAAGTCTCGGCATCTCCACCGGCGACGTG	240
Db	205	GTCCAAACGCGTGTGTCGGAGCGCTCCGGGGCAAGTCTCGGCATCTCCACCGGCGACGTG	264
Qy	241	ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCCACCGCGATCGCGGACGCGCTTAAC	300
Db	265	ATCACCGCGTTCGACGGCGCTCCGATCAACTCGGCCACCGCGATCGCGGACGCGCTTAAC	324
Qy	301	GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGCGGT	360
Db	325	GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGCGCGGT	384
Qy	361	ACAGGGAACGTTGACATTCGGCCGAGGAGACCCCGGCGC	396
Db	385	ACAGGGAACGTTGACATTCGGCCGAGGAGACCCCGGCGC	420
RESULT 24			
AD47110			
ID	AD47110	standard; DNA; 2808 BP.	
XX	AD47110;		
AC	AD47110;		
DT	29-AUG-2003	(revised)	
DT	27-JAN-2003	(first entry)	
XX	Mycobacterium sp. MTB72F-Leishmania sp. MAPS	(aka r95f)-fusion DNA.	
DE	Vaccine; immunity; diagnostic agent; gene therapy; MTB72F; MAPS;		
KW	chimeric; gene; ds.		
XX	Mycobacterium sp.		
OS	Leishmania sp.		
OS	Chimeric.		
XX	Key	Location/Qualifiers	
PH	CDS	4...2796	
FT	FT	/*tag= a	
FT	FT	/product= "Mycobacterium sp. MTB72F-Leishmania sp. MAPS	
FT	FT	(aka r95f) fusion protein"	
XX	WO200272792-A2.		
PN	19-SEP-2002.		
XX	13-MAR-2002; 2002WO-US008223.		
XX	13-MAR-2001; 2001US-0275837P.		
XX	(CORI-) CORIXA CORP.		
PI	Skeiky Y, Brannon M, Guderian J;		
XX	WPI; 2002-759844/82.		
DR	P-PSDB; AAE29731.		
XX	New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,		
PT	M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity		
PT	against pathogenic microorganisms e.g. Leishmania and Mycobacterium		
PT	tuberculosis.		
XX	Example 6; Page 128-129; 155pp; English.		
PS	The invention relates to a recombinant nucleic acid molecule encoding a		
CC	fusion polypeptide. The recombinant nucleic acid comprises a heterologous		
CC	polynucleotide sequence encoding an antigen or an antigenic fragment from		
CC	Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a		
CC	polypeptide or its fragment. The Leishmania polynucleotide is selected		
CC	from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention		



CC are used in methods for eliciting immune response in mammals. They are  
CC useful as vaccines to elicit protective immunity against pathogenic  
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
CC polypeptides are used for enhancing the expression of polynucleotides, as  
CC in vivo diagnostic agents and for raising antibodies in a non-human  
CC animal. The invention is used in gene therapy. The present sequence is  
CC Mycobacterium sp. MTB72F-Leishmania sp. thiol specific antioxidant [TSA;  
CC MAPS (aka r95f)] fusion DNA. This sequence comprises Mycobacterium sp.  
CC MTB72F (a 72 kDa poly-protein fusion construct comprising Ral12-TD9-Ra35)  
CC linked to the Leishmania TSA (MAPS). (Updated on 29-AUG-2003 to  
CC standardise OS field)

XX Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;  
Query Match 100.0%; Score 396; DB 6; Length 2808;  
Best Local Similarity 100.0%; Pred. No. 5.9e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCGCGTCCGATTAATCCAGCTGTCACAGGTCGGCAGGATTCGCGATTCGCGATC 60  
Db |||||||  
25 ACGCCGCGTCCGATTAATCCAGCTGTCACAGGTCGGCAGGATTCGCGATTCGCGATC 84  
Qy 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCAACCGTTTCAT 120  
Db |||||||  
85 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCAACCGTTTCAT 144  
Qy 121 ATCGGGCTACCGCTTCTCGGCTGGTGTGTCACACACACGCGGCGCAGCA 180  
Db |||||||  
145 ATCGGGCTACCGCTTCTCGGCTGGTGTGTCACACACACGCGGCGCAGCA 204  
Qy 181 GTCCAAACCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db |||||||  
205 GTCCAAACCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 264  
Qy 241 ATCCACCGGTCGACGGGCTCCGATCAATCTGGCCACCGGATGGCGAGCGCTTAAC 300  
Db |||||||  
265 ATCCACCGGTCGACGGGCTCCGATCAATCTGGCCACCGGATGGCGAGCGCTTAAC 324  
Qy 301 GGGCATCATCCCGTGAAGTCAATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGACCGGT 360  
Db |||||||  
325 GGGCATCATCCCGTGAAGTCAATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGACCGGT 384  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db |||||||  
385 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 420

RESULT 25  
ADA26357  
ID ADA26357 standard; DNA; 2808 BP.  
AC  
XX  
XX  
DT 20-NOV-2003 (first entry)  
XX Mycobacterium MTB72F-MAPS (fusion r95F) protein encoding DNA.  
DE ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
XX tuberculosis; tuberculostatic; gene therapy; vaccine.  
XX Chimeric.  
OS Mycobacterium sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 4..2796  
FT /\*tag= a  
FT /product= "MTB72F-MAPS (fusion r95F)"  
XX  
XX WO2003070187-A2.  
XX  
XX 28-AUG-2003.  
XX  
XX 18-FEB-2003; 2003WO-US004903.

XX 15-FEB-2002; 2002US-0357351P.  
XX (CORI-) CORIXA CORP.  
XX Skeiky Y, Guderian J, Reed S;  
XX WPI: 2003-697554/66.  
XX P-PSDB; ADA26364.  
XX New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
XX and MTB39 antigens, with or without the MTB85A antigen, from a  
XX Mycobacterium species, useful for diagnosing, preventing and/or treating  
XX tuberculosis infection.  
XX Disclosure, Fig 5; 112pp; English.  
XX The invention relates to a novel nucleic acid encoding a fusion  
XX polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
XX MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
XX A polypeptide of the invention has tuberculostatic activity. A  
XX polynucleotide of the invention may have a use in gene therapy, and as a  
XX vaccine. The methods and compositions of the present invention are useful  
XX for diagnosing, preventing and/or treating tuberculosis infection. The  
XX present sequence is used in the exemplification of the invention.  
XX Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 2808;  
Best Local Similarity 100.0%; Pred. No. 5.9e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGCCGCGTCCGATTAATCCAGCTGTCACAGGTCGGCAGGATTCGCGATTCGCGATC 60  
Db |||||||  
25 ACGCCGCGTCCGATTAATCCAGCTGTCACAGGTCGGCAGGATTCGCGATTCGCGATC 84  
Qy 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCAACCGTTTCAT 120  
Db |||||||  
85 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCAACCGTTTCAT 144  
Qy 121 ATCGGGCTACCGCTTCTCGGCTGGTGTGTCACACACACGCGGCGCAGCA 180  
Db |||||||  
145 ATCGGGCTACCGCTTCTCGGCTGGTGTGTCACACACACGCGGCGCAGCA 204  
Qy 181 GTCCAAACCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db |||||||  
205 GTCCAAACCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 264  
Qy 241 ATCCACCGGTCGACGGGCTCCGATCAATCTGGCCACCGGATGGCGAGCGCTTAAC 300  
Db |||||||  
265 ATCCACCGGTCGACGGGCTCCGATCAATCTGGCCACCGGATGGCGAGCGCTTAAC 324  
Qy 301 GGGCATCATCCCGTGAAGTCAATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGACCGGT 360  
Db |||||||  
325 GGGCATCATCCCGTGAAGTCAATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGACCGGT 384  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db |||||||  
385 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 420

RESULT 26  
ADA26355  
ID ADA26355 standard; DNA; 3030 BP.  
XX  
AC ADA26355;  
XX  
XX 20-NOV-2003 (first entry)  
XX Mycobacterium MTB-102F fusion protein encoding DNA.  
XX ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
XX tuberculosis; tuberculostatic; gene therapy; vaccine.  
KW



[illegible]



XX The invention relates to a novel nucleic acid encoding a fusion  
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
CC A polypeptide of the invention may have a use in gene therapy, and as a  
CC polynucleotide of the invention may have a use in gene therapy, and as a  
CC vaccine. The methods and compositions of the present invention are useful  
CC for diagnosing, preventing and/or treating tuberculosis infection. The  
CC present sequence is used in the exemplification of the invention.

XX SQ Sequence 3474 BP; 548 A; 1131 C; 1252 G; 543 T; 0 U; 0 Other;

Query Match 100.0%; Score 396; DB 8; Length 3474;  
Best Local Similarity 100.0%; Pred. No. 6e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAATTCAGCTGCCAGGTGGGAGGATTCGCATTCGGATC 60  
DB 25 ACGGCGCGTCCGATTAATTCAGCTGCCAGGTGGGAGGATTCGCATTCGGATC 84  
QY 61 GGGCAGCGATGGGATCGCGGCGCAGATCCGATCGGTGGGGGTCAACACCGTTTCAT 120  
DB 85 GGGCAGCGATGGGATCGCGGCGCAGATCCGATCGGTGGGGGTCAACACCGTTTCAT 144  
QY 121 ATCGGGCTACCGCTTCTCGGTGGGTGTTGTCGACAAACGCGCAGCA 180  
DB 145 ATCGGGCTACCGCTTCTCGGTGGGTGTTGTCGACAAACGCGCAGCA 204  
QY 181 GTCCAAACGCGTGGGAGGCGTCCGCGCAAGTCTCGCATCTCCACCGGCGAGTG 240  
DB 205 GTCCAAACGCGTGGGAGGCGTCCGCGCAAGTCTCGCATCTCCACCGGCGAGTG 264  
QY 241 ATCCAGCGGTGGGAGGCGTCCGATCAACTCGGCGCAGCGATGGGAGCGCTTAAC 300  
DB 265 ATCCAGCGGTGGGAGGCGTCCGATCAACTCGGCGCAGCGATGGGAGCGCTTAAC 324  
QY 301 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGCGGCGCAGCGT 360  
DB 325 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGCGGCGCAGCGT 384  
QY 361 ACAGGAAACGTGACATTTGGCGAGGACCCCGGCC 396  
DB 385 ACAGGAAACGTGACATTTGGCGAGGACCCCGGCC 420

RESULT 30  
AAI99682\_01

Continuation (2 of 45) of AAI99682 from base 100001 (Mycobacterium tuberculosis strain H  
WP Sequence split into 45 fragments Locus AAI99682 Accession AAI99682

Fragment Name	Begin	End
WP AAI99682_00	1	110000
WP AAI99682_01	100001	210000
WP AAI99682_02	200001	310000
WP AAI99682_03	300001	410000
WP AAI99682_04	400001	510000
WP AAI99682_05	500001	610000
WP AAI99682_06	600001	710000
WP AAI99682_07	700001	810000
WP AAI99682_08	800001	910000
WP AAI99682_09	900001	1010000
WP AAI99682_10	1000001	1110000
WP AAI99682_11	1100001	1210000
WP AAI99682_12	1200001	1310000
WP AAI99682_13	1300001	1410000
WP AAI99682_14	1400001	1510000
WP AAI99682_15	1500001	1610000
WP AAI99682_16	1600001	1710000
WP AAI99682_17	1700001	1810000
WP AAI99682_18	1800001	1910000
WP AAI99682_19	1900001	2010000
WP AAI99682_20	2000001	2110000
WP AAI99682_21	2100001	2210000
WP AAI99682_22	2200001	2310000

WP AAI99682\_23 2300001 2410000  
WP AAI99682\_24 2400001 2510000  
WP AAI99682\_25 2500001 2610000  
WP AAI99682\_26 2600001 2710000  
WP AAI99682\_27 2700001 2810000  
WP AAI99682\_28 2800001 2910000  
WP AAI99682\_29 2900001 3010000  
WP AAI99682\_30 3000001 3110000  
WP AAI99682\_31 3100001 3210000  
WP AAI99682\_32 3200001 3310000  
WP AAI99682\_33 3300001 3410000  
WP AAI99682\_34 3400001 3510000  
WP AAI99682\_35 3500001 3610000  
WP AAI99682\_36 3600001 3710000  
WP AAI99682\_37 3700001 3810000  
WP AAI99682\_38 3800001 3910000  
WP AAI99682\_39 3900001 4010000  
WP AAI99682\_40 4000001 4110000  
WP AAI99682\_41 4100001 4210000  
WP AAI99682\_42 4200001 4310000  
WP AAI99682\_43 4300001 4410000  
WP AAI99682\_44 4400001 4411529

Query Match 100.0%; Score 396; DB 4; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 7.5e-84;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAATTCAGCTGCCAGGTGGGAGGATTCGCATTCGGATC 60  
DB 51815 ACGGCGCGTCCGATTAATTCAGCTGCCAGGTGGGAGGATTCGCATTCGGATC 51874  
QY 61 GGGCAGCGATGGGATCGCGGCGCAGATCCGATCGGTGGGGGTCAACACCGTTTCAT 120  
DB 51875 GGGCAGCGATGGGATCGCGGCGCAGATCCGATCGGTGGGGGTCAACACCGTTTCAT 51934  
QY 121 ATCGGGCTACCGCTTCTCGGTGGGTGTTGTCGACAAACGCGCAGCA 180  
DB 51935 ATCGGGCTACCGCTTCTCGGTGGGTGTTGTCGACAAACGCGCAGCA 51994  
QY 181 GTCCAAACGCGTGGGAGGCGTCCGCGCAAGTCTCGCATCTCCACCGGCGAGTG 240  
DB 51995 GTCCAAACGCGTGGGAGGCGTCCGCGCAAGTCTCGCATCTCCACCGGCGAGTG 52054  
QY 241 ATCCAGCGGTGGGAGGCGTCCGATCAACTCGGCGCAGCGATGGGAGCGCTTAAC 300  
DB 52055 ATCCAGCGGTGGGAGGCGTCCGATCAACTCGGCGCAGCGATGGGAGCGCTTAAC 52114  
QY 301 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGCGGCGCAGCGT 360  
DB 52115 GGGCATCATCCGCTGAGCTCATCTCGGTGACCTGGCAAAACGCGGCGCAGCGT 52174  
QY 361 ACAGGAAACGTGACATTTGGCGAGGACCCCGGCC 396  
DB 52175 ACAGGAAACGTGACATTTGGCGAGGACCCCGGCC 52210

Search completed: April 30, 2004, 05:15:25  
Job time : 297 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2004, 05:10:50 ; Search time 293 Seconds  
(without alignments)

6113.624 Million cell updates/sec

Title: US-09-684-215b-3

Perfect score: 396

Sequence: 1 agggcggtcgataactt.....tggcgaggagccccggcc 396

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	100.0	702	9	US-09-287-849-27
2	396	100.0	702	15	US-10-359-460-27
3	396	100.0	1002	15	US-10-098-732A-3
4	396	100.0	1002	15	US-10-098-732A-5
5	396	100.0	1068	9	US-09-712-363-15
6	396	100.0	2181	16	US-10-369-983-1
7	396	100.0	2190	15	US-10-098-732A-17
8	396	100.0	2287	9	US-09-287-849-1
9	396	100.0	2287	15	US-10-359-460-1
10	396	100.0	2287	15	US-10-098-732A-15
11	396	100.0	2287	16	US-10-359-459-1
12	396	100.0	2451	16	US-10-369-983-8
13	396	100.0	2487	16	US-10-369-983-7
14	396	100.0	2637	16	US-10-369-983-6

15	396	100.0	2808	15	US-10-098-732A-64	Sequence 64, Appl
16	396	100.0	2808	16	US-10-369-983-5	Sequence 5, Appl
17	396	100.0	3030	16	US-10-369-983-3	Sequence 3, Appl
18	396	100.0	3060	16	US-10-369-983-11	Sequence 11, Appl
19	396	100.0	3104	16	US-10-369-983-10	Sequence 10, Appl
20	396	100.0	3474	16	US-10-369-983-9	Sequence 9, Appl
21	394.4	99.6	447	15	US-10-193-003-4	Sequence 4, Appl
22	394.4	99.6	447	15	US-10-084-843-4	Sequence 4, Appl
23	394.4	99.6	447	15	US-10-098-732A-9	Sequence 9, Appl
24	391.2	98.8	1872	15	US-10-193-003-17	Sequence 17, Appl
25	391.2	98.8	1872	15	US-10-084-843-17	Sequence 17, Appl
26	391.2	98.8	1872	15	US-10-098-732A-1	Sequence 1, Appl
27	354	89.4	675	9	US-09-759-143-822	Sequence 822, App
28	354	89.4	675	9	US-09-780-569-822	Sequence 822, App
29	354	89.4	675	9	US-09-822-837-822	Sequence 822, App
30	354	89.4	675	9	US-09-835-793-822	Sequence 822, App
31	354	89.4	675	9	US-09-835-814-822	Sequence 822, App
32	354	89.4	675	14	US-10-012-896-822	Sequence 822, App
33	354	89.4	675	15	US-10-144-678A-822	Sequence 822, App
34	354	89.4	675	15	US-10-294-025-822	Sequence 822, App
35	354	89.4	822	9	US-09-736-457-1862	Sequence 1862, Ap
36	354	89.4	822	9	US-09-902-941-1862	Sequence 1862, Ap
37	354	89.4	822	9	US-09-849-626-1862	Sequence 1862, Ap
38	354	89.4	822	13	US-10-283-017-1862	Sequence 1862, Ap
39	354	89.4	822	15	US-10-017-754-1862	Sequence 1862, Ap
40	354	89.4	822	15	US-10-113-872-1862	Sequence 236, App
41	354	89.4	894	13	US-09-878-722-236	Sequence 236, App
42	354	89.4	894	13	US-09-904-456-236	Sequence 353, App
43	354	89.4	900	9	US-09-735-705-353	Sequence 353, App
44	354	89.4	900	9	US-09-850-716A-353	Sequence 353, App
45	354	89.4	900	9	US-09-897-778-353	Sequence 353, App

#### ALIGNMENTS

#### RESULT 1

US-09-287-849-27 ; Sequence 27, Application US/09287849  
; Patent No. US20020009459A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1998-04-07 /US/818,112  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 702  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion  
; OTHER INFORMATION: protein Ra12-DPPD (designated Mt24), reading  
; OTHER INFORMATION: frame 1  
; NAME/KEY: CDS



```

; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Ra35 mature
US-10-098-732A-3

Query Match
Best Local Similarity 100.0%; Score 396; DB 15; Length 1002;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCCGGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60
DB 598 ACGGCCGGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 657
QY 61 GGGCAGGGGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 120
DB 658 GGGCAGGGGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 717
QY 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACCAACGCGCAACGGCGACGA 180
DB 718 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACCAACGCGCAACGGCGACGA 777
QY 181 GTCCAAACGGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGACGTG 240
DB 778 GTCCAAACGGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGACGTG 837
QY 241 ATCACCGGGTGCAGCGGCTCCGATCAATCTCGGCCACCGCGATGGCGGACGGTTAAC 300
DB 838 ATCACCGGGTGCAGCGGCTCCGATCAATCTCGGCCACCGCGATGGCGGACGGTTAAC 897
QY 301 GGGCATCATCCGGTGAGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGACGGGT 360
DB 898 GGGCATCATCCGGTGAGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGACGGGT 957
QY 361 ACAGGGAACGTGACATTTGCCCGAGGACCCCGGCC 396
DB 958 ACAGGGAACGTGACATTTGCCCGAGGACCCCGGCC 993

RESULT 5
US-09-712-363-15
; Sequence 15, Application US/09712363
; Patent No. US2002016458A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-15

Query Match
Best Local Similarity 100.0%; Score 396; DB 9; Length 1068;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCCGGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60
DB 598 ACGGCCGGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 657
QY 61 GGGCAGGGGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 120
DB 658 GGGCAGGGGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 717
QY 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACCAACGCGCAACGGCGACGA 180
DB 718 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTGTCGACCAACGCGCAACGGCGACGA 777
QY 181 GTCCAAACGGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGACGTG 240
DB 778 GTCCAAACGGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGGACGTG 837
QY 241 ATCACCGGGTGCAGCGGCTCCGATCAATCTCGGCCACCGCGATGGCGGACGGTTAAC 300
DB 838 ATCACCGGGTGCAGCGGCTCCGATCAATCTCGGCCACCGCGATGGCGGACGGTTAAC 897
QY 301 GGGCATCATCCGGTGAGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGACGGGT 360
DB 898 GGGCATCATCCGGTGAGTCTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGACGGGT 957
QY 361 ACAGGGAACGTGACATTTGCCCGAGGACCCCGGCC 396
DB 958 ACAGGGAACGTGACATTTGCCCGAGGACCCCGGCC 993

RESULT 4
US-10-098-732A-5
; Sequence 5, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; Heterologous Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Ra35FLMutSa
US-10-098-732A-5

Query Match
Best Local Similarity 100.0%; Score 396; DB 15; Length 1002;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 670 ACGGCGCGTCCGATTAATCTCCAGTGTCCAGAGGTGGGAGGATTTCGGCATTCGATC 729  
Qy 61 GGGCAGCGGATGGCGATCCGCGGCGAGATCCGATCGGTTGGGGTCAACCCACCGTTTCAT 120  
Db 730 GGGCAGCGGATGGCGATCCGCGGCGAGATCCGATCGGTTGGGGTCAACCCACCGTTTCAT 789  
Qy 121 ATCGGCGCTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACAAACGCGGCGACGA 180  
Db 790 ATCGGCGCTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACAAACGCGGCGACGA 849  
Qy 181 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCGAGTCTCGGCAATCTCCACCGGCGAGTG 240  
Db 850 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCGAGTCTCGGCAATCTCCACCGGCGAGTG 909  
Qy 241 ATCACCGGCTGACGCGCTCCGATCAACTCGGCGGCGAGTGGGAGCGCGTTAAC 300  
Db 910 ATCACCGGCTGACGCGCTCCGATCAACTCGGCGGCGAGTGGGAGCGCGTTAAC 969  
Qy 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGCACGCT 360  
Db 970 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGCACGCT 1029  
Qy 361 ACAGGAACTGACATGTCGCGGAGGACCCCGGCC 396  
Db 1030 ACAGGAACTGACATGTCGCGGAGGACCCCGGCC 1065

RESULT 6  
US-10-369-983-1  
; Sequence 1, Application US/10369983  
; Publication No. US20030235593A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderian, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; PRIOR FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: Description of Artificial Sequence:mutated  
; OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32MutSA)  
US-10-369-983-1

Query Match 100.0%; Score 396; DB 16; Length 2181;  
Best Local Similarity 100.0%; Pred. No. 3e-107;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGGCGCTCCGATTAATCTCCAGTGTCCAGAGGTGGGAGGATTTCGGCATTCGATC 60  
Db 598 ACGGCGCTCCGATTAATCTCCAGTGTCCAGAGGTGGGAGGATTTCGGCATTCGATC 657  
Qy 61 GGGCAGCGGATGGCGATCCGCGGCGAGATCCGATCGGTTGGGGTCAACCCACCGTTTCAT 120  
Db 658 GGGCAGCGGATGGCGATCCGCGGCGAGATCCGATCGGTTGGGGTCAACCCACCGTTTCAT 717  
Qy 121 ATCGGCGCTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACAAACGCGGCGACGA 180  
Db 718 ATCGGCGCTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACAAACGCGGCGACGA 777  
Qy 181 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCGAGTCTCGGCAATCTCCACCGGCGAGTG 240  
Db 778 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCGAGTCTCGGCAATCTCCACCGGCGAGTG 837

Qy 241 ATCACCGGCTGACGCGCTCCGATCAACTCGGCGGCGAGTGGGAGCGCGTTAAC 300  
Db 838 ATCACCGGCTGACGCGGCTCCGATCAACTCGGCGGCGAGTGGGAGCGCGTTAAC 897  
Qy 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGCACGCT 360  
Db 898 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGCACGCT 957  
Qy 361 ACAGGAACTGACATGTCGCGGAGGACCCCGGCC 396  
Db 958 ACAGGAACTGACATGTCGCGGAGGACCCCGGCC 993

RESULT 7  
US-10-098-732A-17  
; Sequence 17, Application US/10098732A  
; Publication No. US20030175294A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Brannon, Mark  
; APPLICANT: Guderian, Jeffrey  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
; FILE REFERENCE: 014058-012010US  
; CURRENT APPLICATION NUMBER: US/10/098,732A  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR FILING DATE: 2001-03-13  
; PRIOR APPLICATION NUMBER: US 60/275,837  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 2190  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:MTB72PMutSA  
; OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)  
US-10-098-732A-17

Query Match 100.0%; Score 396; DB 15; Length 2190;  
Best Local Similarity 100.0%; Pred. No. 3e-107;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGGCGCGTCCGATTAATCTCCAGTGTCCAGAGGTGGGAGGATTTCGGCATTCGATC 60  
Db 22 ACGGCGCGTCCGATTAATCTCCAGTGTCCAGAGGTGGGAGGATTTCGGCATTCGATC 81  
Qy 61 GGGCAGCGGATGGCGATCCGCGGCGAGATCCGATCGGTTGGGGTCAACCCACCGTTTCAT 120  
Db 82 GGGCAGCGGATGGCGATCCGCGGCGAGATCCGATCGGTTGGGGTCAACCCACCGTTTCAT 141  
Qy 121 ATCGGCGCTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACAAACGCGGCGACGA 180  
Db 142 ATCGGCGCTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACAAACGCGGCGACGA 201  
Qy 181 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCGAGTCTCGGCAATCTCCACCGGCGAGTG 240  
Db 202 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCGAGTCTCGGCAATCTCCACCGGCGAGTG 261  
Qy 241 ATCACCGGCTGACGCGCTCCGATCAACTCGGCGGCGAGTGGGAGCGCGCTTTAAC 300  
Db 262 ATCACCGGCTGACGCGCTCCGATCAACTCGGCGGCGAGTGGGAGCGCGCTTTAAC 321  
Qy 301 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGCACGCT 360  
Db 322 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAAAACCAAGTCGGGCGCACGCT 381  
Qy 361 ACAGGAACTGACATGTCGCGGAGGACCCCGGCC 396  
Db 382 ACAGGAACTGACATGTCGCGGAGGACCCCGGCC 417

RESULT 8  
US-09-287-849-1  
; Sequence 1, Application US/09287849  
; Patent No. US2002009459A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2287  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion  
; OTHER INFORMATION: protein Ral2-TbH9-Ra35 (designated Mtb32-Mtb39  
; OTHER INFORMATION: fusion)  
; NAME/KEY: modified\_base  
; LOCATION: (30)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: modified\_base  
; LOCATION: (33)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: CDS  
; LOCATION: (42)..(2231)  
; NAME/KEY: modified\_base  
; LOCATION: (2270)  
; OTHER INFORMATION: n = g, a, c or t  
US-09-287-849-1  
Query Match 100.0%; Score 396; DB 9; Length 2287;  
Best Local Similarity 100.0%; Pred. No. 3.1e-107;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGGCGCGTCCGATTAACCTTCAGCTGCCAGGTTGGGAGTTCCGATTCGATC 60  
Db 63 ACGGCGCGTCCGATTAACCTTCAGCTGCCAGGTTGGGAGTTCCGATTCGATC 122  
Qy 61 GGGCAGGCGGATGGGATCGCGGGCCAGATCCGATCGGTTGGGGTCCACCCACCGTTTCAT 120  
Db 123 GGGCAGGCGGATGGGATCGCGGGCCAGATCCGATCGGTTGGGGTCCACCCACCGTTTCAT 182  
Qy 121 ATCGGGCTACCGCTTCTCCGCTTGGTGTTCGACACAAACAGGCAACGGCGCACGA 180  
Db 183 ATCGGGCTACCGCTTCTCCGCTTGGTGTTCGACACAAACAGGCAACGGCGCACGA 242  
Qy 181 GTCCAAACCGGTGGTCCGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACGGCGAGTG 240  
Db 243 GTCCAAACCGGTGGTCCGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACGGCGAGTG 302  
Qy 241 ATACCGGGGTCCGAGCGGCTCCGATCAACTCGGCGCACCGGATGGGGAGCGCGCTTAAC 300  
Db 303 ATACCGGGGTCCGAGCGGCTCCGATCAACTCGGCGCACCGGATGGGGAGCGCGCTTAAC 362

Qy 301 GGGCATCATCCCGTGGAGTCACTCGGTGACCTGGCAACCAAGTGGGGCGGACCGGT 360  
Db 363 GGGCATCATCCCGTGGAGTCACTCGGTGACCTGGCAACCAAGTGGGGCGGACCGGT 422  
Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
Db 423 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 458  
RESULT 9  
US-10-359-460-1  
; Sequence 1, Application US/10359460  
; Publication No. US20030147911A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/10/359,460  
; CURRENT FILING DATE: 2003-02-05  
; PRIOR APPLICATION NUMBER: US/09/287,849  
; PRIOR FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2287  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion  
; OTHER INFORMATION: protein Ral2-TbH9-Ra35 (designated Mtb32-Mtb39  
; OTHER INFORMATION: fusion)  
; NAME/KEY: modified\_base  
; LOCATION: (30)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: modified\_base  
; LOCATION: (33)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: CDS  
; LOCATION: (42)..(2231)  
; NAME/KEY: modified\_base  
; LOCATION: (2270)  
; OTHER INFORMATION: n = g, a, c or t  
US-10-359-460-1  
Query Match 100.0%; Score 396; DB 15; Length 2287;  
Best Local Similarity 100.0%; Pred. No. 3.1e-107;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGGCGCGTCCGATTAACCTTCAGCTGCCAGGTTGGGAGTTCCGATTCGATTCGATC 60  
Db 63 ACGGCGCGTCCGATTAACCTTCAGCTGCCAGGTTGGGAGTTCCGATTCGATTCGATC 122



```

RESULT 10
US-10-098-732A-15
; Sequence 15, Application US/10098732A
; Publication NO. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; OTHER INFORMATION: protein MTB72F (Ra12-TbH9-Ra35 or MTB32-MTB39
; OTHER INFORMATION: fusion)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; OTHER INFORMATION: MTB72F
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1)..(2287)
; OTHER INFORMATION: n = g, a, c or t
; US-10-098-732A-15

```

Query Match	100.0%	Score 396	DB 15	Length 2287
Best Local Similarity	100.0%	Pred. No. 3.1e-107		
Matches 396	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ACGGCCGCGTCCGATAACTTCCAGCTGTCCACGGTGGCAGGGAATCGCCATTCGGATC	60	
Db	63	ACGGCCGCGTCCGATAACTTCCAGCTGTCCACGGTGGCAGGGAATCGCCATTCGGATC	122	
QY	61	GGCAGCGCATGGGATTCGGCGGCACAGATCCGATTCGGGTGGGGGTCAACCCACGGTTCAT	120	
Db	123	GGCAGCGCATGGGATTCGGCGGCACAGATCCGATTCGGGTGGGGGTCAACCCACGGTTCAT	182	

```

RESULT 11
US-10-359-459-1
; Sequence 1, Application US/10359459
; Publication No. US20040013677A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/10/359,459
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-RaJ5
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
; US-10-359-459-1

```

	Query Match	100.0%	Score 396;	DB 16;	Length 2287;
	Best Local Similarity	100.0%;	Pred. No. 3.1e-107;		
	Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	ACGGCCGGTCCGATAACTCCAGCTGTCCAGGTGGGAGGATTCGCATTCGCATC	60		
Db	63	ACGGCCGGTCCGATAACTTCAGCTGTCCAGGTGGGAGGATTCGCATTCGCATC	122		
QY	61	GGGAGGCGGATGGCGATCTCGGGGCGAGATCCGATCGGTTGGGGGTGCACCCACCGTTCAT	120		

Db 123 GGGCAGCGATGCGGATCGCGGGCCAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAAT 182  
 Qy 121 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 180  
 Db 183 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 242  
 Qy 181 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 240  
 Db 243 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 302  
 Qy 241 ATCCCGGGTCGACGGGCTCCGATCAACTCGGCCACCGGATCGGGGACGCGCTTAAC 300  
 Db 303 ATCCCGGGTCGACGGGCTCCGATCAACTCGGCCACCGGATCGGGGACGCGCTTAAC 362  
 Qy 301 GGGCATCATCCCGGTGAGCTCATCTCGGTGACCTTGGGCAACCAAGTGGGGGGACCGGT 360  
 Db 363 GGGCATCATCCCGGTGAGCTCATCTCGGTGACCTTGGGCAACCAAGTGGGGGGACCGGT 422  
 Qy 361 ACAGGAACTGACATTCGGCGAGGACCCCGGCC 396  
 Db 423 ACAGGAACTGACATTCGGCGAGGACCCCGGCC 458

RESULT 12  
 US-10-369-983-8  
 ; Sequence 8, Application US/10369983  
 ; Publication No. US20030235593A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Guderian, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-009081US  
 ; CURRENT APPLICATION NUMBER: US/10/369,983  
 ; PRIOR FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 2451  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; MTB81F (MTB72F-DPV)  
 US-10-369-983-8

Query Match 100.0%; Score 396; DB 16; Length 2451;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-107;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGGCGGCTCGGATTAATTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 60  
 Db 25 ACAGGCGGCTCGGATTAATTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 84  
 Qy 61 GGGCAGCGATGCGGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCACCGTTTCAAT 120  
 Db 85 GGGCAGCGATGCGGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCACCGTTTCAAT 144  
 Qy 121 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 180  
 Db 145 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 204  
 Qy 181 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 240  
 Db 85 GGGCAGCGATGCGGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCACCGTTTCAAT 144  
 Qy 121 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 180  
 Db 145 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 204  
 Qy 181 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 240  
 Db 205 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 264  
 Qy 241 ATCCCGGGTCGACGGGCTCCGATCAACTCGGCCACCGGATCGGGGACGCGCTTAAC 300  
 Db 243 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 360  
 Qy 361 ACAGGAACTGACATTCGGCGAGGACCCCGGCC 396  
 Db 423 ACAGGAACTGACATTCGGCGAGGACCCCGGCC 458

RESULT 14  
 US-10-369-983-6  
 ; Sequence 6, Application US/10369983  
 ; Publication No. US20030235593A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Guderian, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-009081US  
 ; CURRENT APPLICATION NUMBER: US/10/369,983  
 ; PRIOR FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 2487  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; MTB72F-MTI  
 US-10-369-983-7

Query Match 100.0%; Score 396; DB 16; Length 2487;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-107;  
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGGCGGCTCGGATTAATTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 60  
 Db 25 ACAGGCGGCTCGGATTAATTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 84  
 Qy 61 GGGCAGCGATGCGGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCACCGTTTCAAT 120  
 Db 85 GGGCAGCGATGCGGATCGCGGGCCAGATCCGATCGGTGGGGGTCAACCACCGTTTCAAT 144  
 Qy 121 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 180  
 Db 145 ATCCGGGCTACCGGCTTCTCTCGGCTTGGGTGTTGTGACAAACACGCGCAACGCGCACGA 204  
 Qy 181 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 240  
 Db 205 GTCCAAACGCGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGCTG 264  
 Qy 241 ATCCCGGGTCGACGGGCTCCGATCAACTCGGCCACCGGATCGGGGACGCGCTTAAC 300  
 Db 265 ATCCCGGGTCGACGGGCTCCGATCAACTCGGCCACCGGATCGGGGACGCGCTTAAC 364  
 Qy 301 GGGCATCATCCCGGTGAGCTCATCTCGGTGACCTTGGGCAACCAAGTGGGGGGACCGGT 360  
 Db 325 GGGCATCATCCCGGTGAGCTCATCTCGGTGACCTTGGGCAACCAAGTGGGGGGACCGGT 384  
 Qy 361 ACAGGAACTGACATTCGGCGAGGACCCCGGCC 396  
 Db 385 ACAGGAACTGACATTCGGCGAGGACCCCGGCC 420

```

; APPLICANT: Skeiky, Yasir
; APPLICANT: Gudarian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: MTB89F (MTB72F-Erd14)
US-10-369-983-5

Query Match
Best Local Similarity 100.0%; Score 396; DB 16; Length 2637;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCTCCAGGCTGGGCGAGGATTCGCAATCCGATC 60
DB 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCCTCCAGGCTGGGCGAGGATTCGCAATCCGATC 84
QY 61 GGGCAGGCGATGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTTCACCCACCGTTTCA 120
DB 85 GGGCAGGCGATGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTTCACCCACCGTTTCA 144
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACAAACGCGCAACGCGCACGA 180
DB 145 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACAAACGCGCAACGCGCACGA 204
QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 240
DB 205 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 264
QY 241 ATCACCGCGTCCGACGCGCTCCGATCAACTCGGCGCACCGCGATGGCGACGCGCTTAAC 300
DB 265 ATCACCGCGTCCGACGCGCTCCGATCAACTCGGCGCACCGCGATGGCGACGCGCTTAAC 324
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGCGCAACCAAGTCGGCGCGCACGCGT 360
DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGCGCAACCAAGTCGGCGCGCACGCGT 384
QY 361 ACAGGGAACGTGACATTCGCGGAGGACCCCGGCGC 396
DB 385 ACAGGGAACGTGACATTCGCGGAGGACCCCGGCGC 420

RESULT 15
US-10-098-732A-64
; Sequence 64, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Gudarian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64

; APPLICANT: Skeiky, Yasir
; APPLICANT: Gudarian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2808
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-5

Query Match
Best Local Similarity 100.0%; Score 396; DB 16; Length 2808;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCTCCAGGCTGGGCGAGGATTCGCAATCCGATC 60
DB 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCCTCCAGGCTGGGCGAGGATTCGCAATCCGATC 84
QY 61 GGGCAGGCGATGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTTCACCCACCGTTTCA 120
DB 85 GGGCAGGCGATGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTTCACCCACCGTTTCA 144
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACAAACGCGCAACGCGCACGA 180
DB 145 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACAAACGCGCAACGCGCACGA 204
QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 240
DB 205 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 264
QY 241 ATCACCGCGTCCGACGCGCTCCGATCAACTCGGCGCACCGCGATGGCGACGCGCTTAAC 300
DB 265 ATCACCGCGTCCGACGCGCTCCGATCAACTCGGCGCACCGCGATGGCGACGCGCTTAAC 324
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGCGCAACCAAGTCGGCGCGCACGCGT 360
DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGCGCAACCAAGTCGGCGCGCACGCGT 384
QY 361 ACAGGGAACGTGACATTCGCGGAGGACCCCGGCGC 396
DB 385 ACAGGGAACGTGACATTCGCGGAGGACCCCGGCGC 420

RESULT 16
US-10-369-983-5
; Sequence 5, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Gudarian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2808
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-5

Query Match
Best Local Similarity 100.0%; Score 396; DB 16; Length 2808;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCCTCCAGGCTGGGCGAGGATTCGCAATCCGATC 60
DB 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCCTCCAGGCTGGGCGAGGATTCGCAATCCGATC 84
QY 61 GGGCAGGCGATGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTTCACCCACCGTTTCA 120
DB 85 GGGCAGGCGATGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTTCACCCACCGTTTCA 144
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACAAACGCGCAACGCGCACGA 180
DB 145 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACAAACGCGCAACGCGCACGA 204
QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 240
DB 205 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 264
QY 241 ATCACCGCGTCCGACGCGCTCCGATCAACTCGGCGCACCGCGATGGCGACGCGCTTAAC 300
DB 265 ATCACCGCGTCCGACGCGCTCCGATCAACTCGGCGCACCGCGATGGCGACGCGCTTAAC 324
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGCGCAACCAAGTCGGCGCGCACGCGT 360
DB 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGCGCAACCAAGTCGGCGCGCACGCGT 384
QY 361 ACAGGGAACGTGACATTCGCGGAGGACCCCGGCGC 396
DB 385 ACAGGGAACGTGACATTCGCGGAGGACCCCGGCGC 420
```

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCGGATC 60  
DB 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCGGATC 84  
QY 61 GGGCAGCGATGCGGATTCGGGCGGATCCGATCGGCTGGGCGGTCACCCACCGTTTAT 120  
DB 85 GGGCAGCGATGCGGATTCGGGCGGATCCGATCGGCTGGGCGGTCACCCACCGTTTAT 144  
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 180  
DB 145 ATCGGCGCTACCGCTTCCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 204  
QY 181 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCAAACGCTG 240  
DB 205 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCAAACGCTG 264  
QY 241 ATACCGCGGTGACGCGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 300  
DB 265 ATACCGCGGTGACGCGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 324  
QY 301 GGGCATCATCCGCTGACGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 360  
DB 325 GGGCATCATCCGCTGACGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 384  
QY 361 ACAGGAAACGCTGACATTCGGCGGAGGACCCCGGCC 396  
DB 385 ACAGGAAACGCTGACATTCGGCGGAGGACCCCGGCC 420

RESULT 17  
US-10-369-983-3  
; Sequence 3, Application US/10369983  
; Publication No. US20030235593A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderian, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; CURRENT FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 3030  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion  
; OTHER INFORMATION: protein  
US-10-369-983-3

Query Match 100.0%; Score 396; DB 16; Length 3030;  
Best Local Similarity 100.0%; Pred. No. 3.1e-107;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCGGATC 60  
DB 595 ACGGCGCGTCCGATTAACCTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCGGATC 654  
QY 61 GGGCAGCGATGCGGATTCGGGCGGATCCGATCGGCTGGGCGGTCACCCACCGTTTAT 120  
DB 655 GGGCAGCGATGCGGATTCGGGCGGATCCGATCGGCTGGGCGGTCACCCACCGTTTAT 714  
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 180  
DB 715 ATCGGCGCTACCGCTTCCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 774  
QY 181 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCAAACGCTG 240

DB 775 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCAAACGCTG 834  
QY 241 ATACCGCGGTGACGCGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 300  
DB 835 ATACCGCGGTGACGCGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 894  
QY 301 GGGCATCATCCGCTGACGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 360  
DB 895 GGGCATCATCCGCTGACGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 954  
QY 361 ACAGGAAACGCTGACATTCGGCGGAGGACCCCGGCC 396  
DB 955 ACAGGAAACGCTGACATTCGGCGGAGGACCCCGGCC 990

RESULT 18  
US-10-369-983-11  
; Sequence 11, Application US/10369983  
; Publication No. US20030235593A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderian, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; CURRENT FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 3060  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein  
; OTHER INFORMATION: MTB103F (MTB72F-85b)  
US-10-369-983-11

Query Match 100.0%; Score 396; DB 16; Length 3060;  
Best Local Similarity 100.0%; Pred. No. 3.1e-107;  
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATTAACCTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCGGATC 60  
DB 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCGGATC 84  
QY 61 GGGCAGCGATGCGGATTCGGGCGGATCCGATCGGCTGGGCGGTCACCCACCGTTTAT 120  
DB 85 GGGCAGCGATGCGGATTCGGGCGGATCCGATCGGCTGGGCGGTCACCCACCGTTTAT 144  
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 180  
DB 145 ATCGGCGCTACCGCTTCCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 204  
QY 181 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCAAACGCTG 240  
DB 205 GTCCAAACGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCAAACGCTG 264  
QY 241 ATACCGCGGTGACGCGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 300  
DB 265 ATACCGCGGTGACGCGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 324  
QY 301 GGGCATCATCCGCTGACGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 360  
DB 325 GGGCATCATCCGCTGACGCTTCGATCACTCGGCGGCAAGTCTCGGATCTCCAAACGCTG 384  
QY 361 ACAGGAAACGCTGACATTCGGCGGAGGACCCCGGCC 396  
DB 385 ACAGGAAACGCTGACATTCGGCGGAGGACCCCGGCC 420

```

RESULT 19
US-10-369-983-10
; Sequence 10, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderman, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3104
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: MTB102tm2F (MTB102FTM, MTB72F-hTCC#1)
US-10-369-983-10

Query Match 100.0%; Score 396; DB 16; Length 3104;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCCGATC 60
Ddb 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCCGATC 84
QY 61 GGGCAGGCGATGGCGATTCGGCGGCAGATCCGATCGGGTGGGGGTCAACACCGTTCAT 120
Ddb 85 GGGCAGGCGATGGCGATTCGGCGGCAGATCCGATCGGGTGGGGGTCAACACCGTTCAT 144
QY 121 ATCGGGCTACCGGCTTCCTCGGCTTGGGTGTGTGCACAAACAGGCAAGGGCGACGA 180
Ddb 145 ATCGGGCTACCGGCTTCCTCGGCTTGGGTGTGTGCACAAACAGGCAAGGGCGACGA 204
QY 181 GTCCAAACGCTGTGCGGAGCGCTTCGATCAACTCGGCCACCGCATCGCGGACGCGTTAAC 240
Ddb 205 GTCCAAACGCTGTGCGGAGCGCTTCGATCAACTCGGCCACCGCATCTCCACCGCGACGTG 264
QY 241 ATCACCAGGCTGCGACGCGCTTCGATCAACTCGGCCACCGCATCGCGGACGCGTTAAC 300
Ddb 265 ATCACCAGGCTGCGACGCGCTTCGATCAACTCGGCCACCGCATCGCGGACGCGTTAAC 324
QY 301 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACACAGTTCGGGCGGACGCGT 360
Ddb 325 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACACAGTTCGGGCGGACGCGT 384
QY 361 ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC 396
Ddb 395 ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC 420

RESULT 20
US-10-369-983-9
; Sequence 9, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderman, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/193.002  
FILING DATE: 10-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072.596  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 210121.417C9  
REFERENCE/DOCKET NUMBER:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-193-002-4  
Query Match 99.6%; Score 394.4; DB 15; Length 447;  
Best Local Similarity 99.7%; Pred. No. 8.8e-107;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATTAACCTCCAGGTGCGGAGGATTCGGCATTCGGATC 60  
DB 11 ACGGCGCGTCCGATTAACCTCCAGGTGCGGAGGATTCGGCATTCGGATC 70  
QY 61 GGGCAGGCGATGCGGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTAT 120  
DB 71 GGGCAGGCGATGCGGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTAT 130  
QY 121 ATCGGGCTACCGCTTCTCGGTTGGGTTGTGACAAACAGCGCAACGGCGACGA 180  
DB 131 ATCGGGCTACCGCTTCTCGGTTGGGTTGTGACAAACAGCGCAACGGCGACGA 190  
QY 181 GTCCAAACGCGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGACGTG 240  
DB 191 GTCCAAACGCGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGACGTG 250  
QY 241 ATCCCGCGGTGACGCGCTCCGATCAATCGCGCAACCGCGATGGCGGCGGCTTAA 300  
DB 251 ATCCCGCGGTGACGCGCTCCGATCAATCGCGCAACCGCGATGGCGGCGGCTTAA 310  
QY 301 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGACGCT 360  
DB 311 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGACGCT 370  
QY 361 ACAGGAACTGACATTCGCGGAGGACCCCGGCC 396  
DB 371 ACAGGAACTGACATTCGCGGAGGACCCCGGCC 406

RESULT 22  
US-10-084-843-4  
Sequence 4, Application US/10084843  
Publication No. US20030143243A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
Skeiky, Yasir A.W.  
Dillon, David C.  
Campos-Neto, Antonio  
Houghton, Raymond  
Vedvick, Thomas S.  
Twardzik, Daniel R.  
Lodes, Michael J.  
Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/084.843  
FILING DATE: 25-Feb-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072.967  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-084-843-4  
Query Match 99.6%; Score 394.4; DB 15; Length 447;  
Best Local Similarity 99.7%; Pred. No. 8.8e-107;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATTAACCTCCAGGTGCGGAGGATTCGGCATTCGGATC 60  
DB 11 ACGGCGCGTCCGATTAACCTCCAGGTGCGGAGGATTCGGCATTCGGATC 70  
QY 61 GGGCAGGCGATGCGGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTAT 120  
DB 71 GGGCAGGCGATGCGGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTAT 130  
QY 121 ATCGGGCTACCGCTTCTCGGTTGGGTTGTGACAAACAGCGCAACGGCGACGA 180  
DB 131 ATCGGGCTACCGCTTCTCGGTTGGGTTGTGACAAACAGCGCAACGGCGACGA 190  
QY 181 GTCCAAACGCGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGACGTG 240  
DB 191 GTCCAAACGCGTGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGACGTG 250  
QY 241 ATCCCGCGGTGACGCGCTCCGATCAATCGCGCAACCGCGATGGCGGCGGCTTAA 300  
DB 251 ATCCCGCGGTGACGCGCTCCGATCAATCGCGCAACCGCGATGGCGGCGGCTTAA 310  
QY 301 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGACGCT 360  
DB 311 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTGGCGGCGACGCT 370  
QY 361 ACAGGAACTGACATTCGCGGAGGACCCCGGCC 396  
DB 371 ACAGGAACTGACATTCGCGGAGGACCCCGGCC 406

RESULT 23  
US-10-098-732A-9  
Sequence 9, Application US/10098732A

Publication No. US20030175294A1  
GENERAL INFORMATION:  
APPLICANT: Skeiky, Yasir  
APPLICANT: Brannon, Mark  
APPLICANT: Guderian, Jeffrey  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
TITLE OF INVENTION: Leishmania Antigen  
FILE REFERENCE: 014058-012010US  
CURRENT APPLICATION NUMBER: US/10/098,732A  
CURRENT FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: 2003-04-29  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 9  
LENGTH: 447  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: MTB32A (MTB32A)  
OTHER INFORMATION: C-terminus)  
US-10-098-732A-9

Query Match 99.6%; Score 394.4; DB 15; Length 447;  
Best Local Similarity 99.7%; Pred. No. 8.8e-107;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCACAGTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60  
DB 11 ACGGCGCGTCCGATAAATTCACAGTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 61 GGGCAGGCGATCGCGATCGCGGCGAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAAT 120  
DB 71 GGGCAGGCGATCGCGATCGCGGCGAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAAT 130  
QY 121 ATCGGGCGTACCGCTTCCGCTTCCGCTTGGTGTTCGACAAACGCGCAACGCGGACGA 180  
DB 131 ATCGGGCGTACCGCTTCCGCTTCCGCTTGGTGTTCGACAAACGCGCAACGCGGACGA 190  
QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
DB 191 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
QY 241 ATCAGCGCGTCAAGCGGCTCCGATCAACTCGGCGACCGGATCGCGGACGCGCTTAAC 300  
DB 251 ATCAGCGCGTCAAGCGGCTCCGATCAACTCGGCGACCGGATCGCGGACGCGCTTAAC 310  
QY 301 GGGCATCATCCCGTGAAGTCACTCGGTGACCTCGGCAACCAAGTCCGGCGGCGACGCGT 360  
DB 311 GGGCATCATCCCGTGAAGTCACTCGGTGACCTCGGCAACCAAGTCCGGCGGCGACGCGT 370  
QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
DB 371 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 406

RESULT 24  
US-10-193-002-17  
Sequence 17, Application US/10193002  
Publication No. US20030135026A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, David C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
TUBERCULOSIS

NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/193,002  
FILING DATE: 10-Jul-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-10-193-002-17

Query Match 98.8%; Score 391.2; DB 15; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.9e-106;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCACAGTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 60  
DB 758 ACGGCGCGTCCGATAAATTCACAGTGTCCAGGGTGGCGAGGATTCGCCATTCGGATC 817  
QY 61 GGGCAGGCGATCGCGATCGCGGCGAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAAT 120  
DB 818 GGGCAGGCGATCGCGATCGCGGCGAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAAT 877  
QY 121 ATCGGGCGTACCGCTTCCGCTTGGTGTTCGACAAACGCGCAACGCGGACGCGT 180  
DB 878 ATCGGGCGTACCGCTTCCGCTTGGTGTTCGACAAACGCGCAACGCGGACGCGT 937  
QY 181 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
DB 938 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997  
QY 241 ATCAGCGCGTCAAGCGGCTCCGATCAACTCGGCGACCGGATCGCGGACGCGCTTAAC 300  
DB 998 ATCAGCGCGTCAAGCGGCTCCGATCAACTCGGCGACCGGATCGCGGACGCGCTTAAC 1057  
QY 301 GGGCATCATCCCGTGAAGTCACTCGGTGACCTCGGCAACCAAGTCCGGCGGCGACGCGT 360  
DB 1058 GGGCATCATCCCGTGAAGTCACTCGGTGACCTCGGCAACCAAGTCCGGCGGCGACGCGT 1117  
QY 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
DB 1118 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 1153

RESULT 25  
US-10-084-843-17  
Sequence 17, Application US/10084843  
Publication No. US20030143243A1



GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
Skeiky, Yasir A.W.  
Dillon, Davin C.  
Campos-Neto, Antonio  
Houghton, Raymond  
Twardzik, Daniel R.  
Lodes, Michael J.  
Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/084,843  
FILING DATE: 25-FEB-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-10-084-843-17  
Query Match 98.8%; Score 391.2; DB 15; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.9e-106;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 60  
DB 758 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 817  
QY 61 GGGCAGGCGATGCGGATCCGAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 120  
DB 818 GGGCAGGCGATGCGGATCCGAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 877  
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 180  
DB 878 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 937  
QY 181 GTCCAAACGGGTGTGGGAGCGCTCCGCGGCGATCCGATCCGATCCGATCCGATCCGAT 240  
DB 938 GTCCAAACGGGTGTGGGAGCGCTCCGCGGCGATCCGATCCGATCCGATCCGATCCGAT 997  
QY 241 ATCAACCGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 300  
DB 998 ATCAACCGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 1057  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGACCTGCAACCAAGTCGGGCGGCGCGT 360  
DB 1118 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGACCTGCAACCAAGTCGGGCGGCGCGT 1117  
RESULT 26  
US-10-098-732A-1  
; Sequence 1, Application US/10098732A  
; Publication No. US20030175294A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Brannon, Mark  
; APPLICANT: Cuderian, Jeffrey  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
; TITLE OF INVENTION: Leishmania Antigen  
; FILE REFERENCE: 014058-012010US  
; CURRENT APPLICATION NUMBER: US/10/098,732A  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: US 60/275,837  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1872  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: MTB32A (Ra35FL)  
; FEATURE:  
; NAME/KEY: modified base  
; LOCATION: (1)..(1872)  
; OTHER INFORMATION: n = g, a, c or t  
US-10-098-732A-1  
Query Match 98.8%; Score 391.2; DB 15; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.9e-106;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 60  
DB 758 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 817  
QY 61 GGGCAGGCGATGCGGATCCGAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 120  
DB 818 GGGCAGGCGATGCGGATCCGAGCTGTCCAGGGTGGGCGAGGATTCGCCATTCCGATC 877  
QY 121 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 180  
DB 878 ATCGGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 937  
QY 181 GTCCAAACGGGTGTGGGAGCGCTCCGCGGCGATCCGATCCGATCCGATCCGATCCGAT 240  
DB 938 GTCCAAACGGGTGTGGGAGCGCTCCGCGGCGATCCGATCCGATCCGATCCGATCCGAT 997  
QY 241 ATCAACCGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 300  
DB 998 ATCAACCGGCTTACCGCTTCTCGGCTTGGGTGTTCGACAAACCGGCAACCGGCGACGA 1057  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGACCTGCAACCAAGTCGGGCGGCGCGT 360  
DB 1118 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGACCTGCAACCAAGTCGGGCGGCGCGT 1117  
RESULT 27  
US-09-759-143-822



RESULT 28  
US-09-780-669-822  
; Sequence 822, Application US/09780669  
; Patent No. US2002065197A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, JIangchun  
; APPLICANT: Dillion, Davin C.

RESULT 29  
US-09-822-827-822  
; Sequence 822, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C1

; CURRENT APPLICATION NUMBER: US/09/822,827  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-827-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.5e-95;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
  
QY 1 ACAGGCGGTCGATACCTTCCAGCTGTCCAGGGTGGGAGGGATTCCGCAATTCGATC 60  
DB 22 ACAGGCGGTCGATACCTTCCAGCTGTCCAGGGTGGGAGGGATTCCGCAATTCGATC 81  
  
QY 61 GGGCAGCGATGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTTCAT 120  
DB 82 GGGCAGCGATGCGATCGCGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
  
QY 121 ATCGGGCTACCGCTTCTCGCTTGGGTGTTGTGACAAACGCGCAACGCGCACGA 180  
DB 130 ATCGGGCTACCGCTTCTCGCTTGGGTGTTGTGACAAACGCGCAACGCGCACGA 189  
  
QY 181 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 240  
DB 190 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 249  
  
QY 241 ATCAACCGCGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAAC 300  
DB 250 ATCAACCGCGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAAC 309  
  
QY 301 GGGCATCATCCGGTGAAGTCACTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 360  
DB 310 GGGCATCATCCGGTGAAGTCACTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 369  
  
QY 361 ACAGGAAACGTGACATTCGCGAGGACCCCGGCC 396  
DB 370 ACAGGAAACGTGACATTCGCGAGGACCCCGGCC 405

## RESULT 30

US-09-895-793-822  
; Sequence 822, Application US/09895793  
; Publication No. US20020192763A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Devin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stoik, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yassir A.W.  
; APPLICANT: Repler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Basols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-895-793-822

Query Match 89.4%; Score 354; DB 9; Length 675;  
Best Local Similarity 95.7%; Pred. No. 7.5e-95;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
  
QY 1 ACAGGCGGTCGATACCTTCCAGCTGTCCAGGGTGGGAGGGATTCCGCAATTCGATC 60  
DB 22 ACAGGCGGTCGATACCTTCCAGCTGTCCAGGGTGGGAGGGATTCCGCAATTCGATC 81  
  
QY 61 GGGCAGCGATGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTTCAT 120  
DB 82 GGGCAGCGATGCGATCGCGGCCAGATC-----AAGCTTCCACCGTTTCAT 129  
  
QY 121 ATCGGGCTACCGCTTCTCGCTTGGGTGTTGTGACAAACGCGCAACGCGCACGA 180  
DB 130 ATCGGGCTACCGCTTCTCGCTTGGGTGTTGTGACAAACGCGCAACGCGCACGA 189  
  
QY 181 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 240  
DB 190 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 249  
  
QY 241 ATCAACCGCGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAAC 300  
DB 250 ATCAACCGCGTCCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGTTAAC 309  
  
QY 301 GGGCATCATCCGGTGAAGTCACTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 360  
DB 310 GGGCATCATCCGGTGAAGTCACTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 369  
  
QY 361 ACAGGAAACGTGACATTCGCGAGGACCCCGGCC 396  
DB 370 ACAGGAAACGTGACATTCGCGAGGACCCCGGCC 405

Search completed: April 30, 2004, 07:45:33  
Job time : 295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2004, 05:13:09 ; Search time 64 Seconds  
(without alignments)  
3433.759 Million cell updates/sec

Title: US-09-684-215B-3

Perfect score: 396  
Sequence: 1 agccgcgcgcgcataactt.....tgccgaggagcccccggcc 396

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCUTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	702	4	US-09-287-849-27
2	396	100.0	2287	4	US-09-223-040-1
3	396	100.0	2287	4	US-09-287-849-1
4	396	100.0	4403765	3	US-09-103-840A-2
5	396	100.0	4411529	3	US-09-103-840A-1
6	394.4	99.6	447	3	US-08-818-112-4
7	394.4	99.6	447	4	US-08-818-111-4
8	394.4	99.6	447	4	US-09-056-556-4
9	394.4	99.6	447	4	US-09-072-596-4
10	394.4	99.6	447	4	US-09-072-967-4
11	391.2	98.8	1872	3	US-08-818-112-17
12	391.2	98.8	1872	4	US-08-818-111-17
13	391.2	98.8	1872	4	US-09-056-556-17
14	391.2	98.8	1872	4	US-09-072-596-17
15	391.2	98.8	1872	4	US-09-072-967-17
16	354	89.4	675	4	US-09-636-215-822
17	354	89.4	675	4	US-09-685-166A-822
18	354	89.4	822	4	US-09-736-457-1862
19	354	89.4	900	4	US-09-643-597-353
20	354	89.4	900	4	US-09-606-421B-353
21	354	89.4	915	4	US-09-636-215-834
22	354	89.4	915	4	US-09-685-166A-834
23	354	89.4	945	4	US-09-736-457-1861
24	354	89.4	1012	4	US-09-643-597-351
25	354	89.4	1012	4	US-09-606-421B-351
26	354	89.4	1203	4	US-09-636-215-851
27	354	89.4	1203	4	US-09-685-166A-851

#### ALIGNMENTS

##### RESULT 1

US-09-287-849-27  
; Sequence 27, Application US/09287849  
; Patent No. 6627198

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Alderson, Mark

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; FILE REFERENCE: 014058-009020US

; CURRENT APPLICATION NUMBER: US/09/287,849

; PRIOR FILING DATE: 1999-04-07

; PRIOR APPLICATION NUMBER: US 08/818,112

; PRIOR FILING DATE: 1997-03-13

; PRIOR APPLICATION NUMBER: US 08/942,578

; PRIOR FILING DATE: 1997-10-01

; PRIOR APPLICATION NUMBER: US 09/025,197

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 09/056,556

; PRIOR FILING DATE: 1998-04-07

; PRIOR APPLICATION NUMBER: US 09/223,040

; PRIOR FILING DATE: 1998-12-30

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 27

; LENGTH: 702

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion

; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading

; OTHER INFORMATION: frame 1

; NAME/KEY: CDS

; LOCATION: (1)..(693)

; OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),

; OTHER INFORMATION: reading frame 1

; NAME/KEY: CDS

; LOCATION: (2)..(700)

; OTHER INFORMATION: reading frame 2

; NAME/KEY: CDS

; LOCATION: (3)..(701)

; OTHER INFORMATION: reading frame 3

; US-09-287-849-27

Query Match

Best Local Similarity 100.0%; Score 396; DB 4; Length 702;

100.0%; Pred. No. 3.2e-95;

Sequence 348, App  
Sequence 348, App  
Sequence 332, App  
Sequence 332, App  
Sequence 188, App  
Sequence 188, App  
Sequence 188, App  
Sequence 188, App  
Sequence 352, App  
Sequence 352, App  
Sequence 336, App  
Sequence 336, App  
Sequence 308, App  
Sequence 308, App  
Sequence 324, App  
Sequence 324, App  
Sequence 316, App  
Sequence 316, App  
Sequence 340, App

28 89.4 1464 4 US-09-620-412C-348  
29 89.4 1464 4 US-09-598-419-348  
30 89.4 1557 4 US-09-620-412C-332  
31 89.4 1557 4 US-09-598-419-332  
32 89.4 1578 4 US-09-556-877-188  
33 89.4 1578 4 US-09-620-412C-188  
34 89.4 1578 4 US-09-598-419-188  
35 89.4 1752 4 US-09-620-412C-352  
36 89.4 1752 4 US-09-598-419-352  
37 89.4 1758 4 US-09-620-412C-336  
38 89.4 1758 4 US-09-598-419-336  
39 89.4 1860 4 US-09-620-412C-308  
40 89.4 1860 4 US-09-598-419-308  
41 89.4 1896 4 US-09-620-412C-324  
42 89.4 1896 4 US-09-598-419-324  
43 89.4 1941 4 US-09-620-412C-316  
44 89.4 1941 4 US-09-598-419-316  
45 89.4 1965 4 US-09-620-412C-340

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCCGCTCCGATAAATTCAGCTGTCAGAGGTGGGAGGATTCGCATTCGGATC 60

Db 25 ACGCCCGCTCCGATAAATTCAGCTGTCAGAGGTGGGAGGATTCGCATTCGGATC 84

Qy 61 GGGCAGCGATGGGATCGGCGGCGAGATCCGATCGGCTGGGGGTCAACCCAGCTTCAT 120

Db 85 GGGCAGCGATGGGATCGGCGGCGAGATCCGATCGGCTGGGGGTCAACCCAGCTTCAT 144

Qy 121 ATCGGGCTACCGCTTCCTCGGCTTCGGTGTGTTGTCGACAAACGCGGCGACGA 180

Db 145 ATCGGGCTACCGCTTCCTCGGCTTCGGTGTGTTGTCGACAAACGCGGCGACGA 204

Qy 181 GTCCAAACGCTGGTGGGAGGCTCCGGCGGCAAGTCTCGCATCTCCAGCGGCGAGTG 240

Db 205 GTCCAAACGCTGGTGGGAGGCTCCGGCGGCAAGTCTCGCATCTCCAGCGGCGAGTG 264

Qy 241 ATCACCAGGCTCGAGCGGCTCCGATCAACTCGGCGCAAGTCTCGCATCTCCAGCGGCGAGTG 300

Db 265 ATCACCAGGCTCGAGCGGCTCCGATCAACTCGGCGCAAGTCTCGCATCTCCAGCGGCGAGTG 324

Qy 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGCGGCGAGCGGT 360

Db 325 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGCGGCGAGCGGT 384

Qy 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCCC 396

Db 385 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCCC 420

RESULT 2

US-09-223-040-1

; Sequence 1, Application US/09223040

; Patent No. 6544522

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; FILE REFERENCE: 014058-009010US

; CURRENT APPLICATION NUMBER: US/09/223,040

; CURRENT FILING DATE: 1998-12-30

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2287

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion

; NAME/KEY: modified\_base

; LOCATION: (30)

; NAME/KEY: modified\_base

; LOCATION: (33)

; NAME/KEY: modified\_base

; LOCATION: (42) ..(2231)

; NAME/KEY: modified\_base

; LOCATION: (2270)

; OTHER INFORMATION: n = g, a, c or t

US-09-223-040-1

Query Match 100.0%; Score 396; DB 4; Length 2287;

Best Local Similarity 100.0%; Pred. No. 4.1e-95;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGCCCGCTCCGATAAATTCAGCTGTCAGAGGTGGGAGGATTCGCATTCGGATC 60

Db 63 ACGCCCGCTCCGATAAATTCAGCTGTCAGAGGTGGGAGGATTCGCATTCGGATC 122

Qy 61 GGGCAGCGATGGGATCGGCGGCGAGATCCGATCGGCTGGGGGTCAACCCAGCTTCAT 120

Db 123 GGGCAGCGATGGGATCGGCGGCGAGATCCGATCGGCTGGGGGTCAACCCAGCTTCAT 182

Qy 121 ATCGGGCTACCGCTTCCTCGGCTTCGGTGTGTTGTCGACAAACGCGGCGACGA 180

Db 183 ATCGGGCTACCGCTTCCTCGGCTTCGGTGTGTTGTCGACAAACGCGGCGACGA 242

Qy 181 GTCCAAACGCTGGTGGGAGGCTCCGGCGGCAAGTCTCGCATCTCCAGCGGCGAGTG 240

Db 243 GTCCAAACGCTGGTGGGAGGCTCCGGCGGCAAGTCTCGCATCTCCAGCGGCGAGTG 302

Qy 241 ATCACCAGGCTCGAGCGGCTCCGATCAACTCGGCGCAAGTTCGGCGGCGAGTG 300

Db 303 ATCACCAGGCTCGAGCGGCTCCGATCAACTCGGCGCAAGTTCGGCGGCGAGTG 362

Qy 301 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGCGGCGAGCGGT 360

Db 363 GGGCATATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTTCGGCGGCGAGCGGT 422

Qy 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCCC 396

Db 423 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCCC 458

RESULT 3

US-09-287-849-1

; Sequence 1, Application US/09287849

; Patent No. 6627198

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Alderson, Mark

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; FILE REFERENCE: 014058-009020US

; CURRENT APPLICATION NUMBER: US/09/287,849

; CURRENT FILING DATE: 1999-04-07

; PRIOR APPLICATION NUMBER: US 08/818,112

; PRIOR FILING DATE: 1997-03-13

; PRIOR APPLICATION NUMBER: US 08/942,578

; PRIOR FILING DATE: 1997-10-01

; PRIOR APPLICATION NUMBER: US 09/025,197

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 09/056,556

; PRIOR FILING DATE: 1998-04-07

; PRIOR APPLICATION NUMBER: US 09/223,040

; PRIOR FILING DATE: 1998-12-30

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2287

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion

; NAME/KEY: modified\_base

; LOCATION: (30)

; NAME/KEY: modified\_base

; LOCATION: (33)

; NAME/KEY: modified\_base

; LOCATION: (42) ..(2231)

; NAME/KEY: modified\_base

; LOCATION: (2270)

; OTHER INFORMATION: n = g, a, c or t

OTHER INFORMATION: n = g, a, c o r t  
US-09-287-849-1

Query Match	Score 396;	DB 4;	Length 2287;
Best Local Similarity	100.0%;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps 0;			

Qy	1	ACGGCCGGCTCCGATAACTTCCAGCTGTCCACAGGGTGGCAGGGATTCGCCATTCCGATC	60
Db	63	ACGGCCGGCTCCGATAACTTCCAGCTGTCCACAGGGTGGCAGGGATTCGCCATTCCGATC	122
Qy	61	GGGCAGCGCATGGCGATCGCGGGCCAGATCCGATCGGTCGGTGGGGGTACCCACCGTTCAT	120
Db	123	GGGCAGCGCATGGCGATCGCGGGCCAGATCCGATCGGTCGGTGGGGGTACCCACCGTTCAT	182
Qy	121	ATCGGGGCTACCGGCTTCTCGGCTTGGGTGTGTGCAACAACAGGCAACGGGCGACGA	180
Db	183	ATCGGGGCTACCGGCTTCTCGGCTTGGGTGTGTGCAACAACAGGCAACGGGCGACGA	242
Qy	181	GTCCAAACGGTGTGGAGAGCTCTCCGGCGGGAAGTCTCGGATCTCCACGGGGAAGTG	240
Db	243	GTCCAAACGGTGTGGAGAGCTCTCCGGCGGGAAGTCTCGGATCTCCACGGGGAAGTG	302
Qy	241	ATCACGGCGTCGAGGGCGCTCCGATCAACTCGGCCACCGCGATCGGCGGACCGGTTAAAC	300
Db	303	ATCACGGCGTCGAGGGCGCTCCGATCAACTCGGCCACCGCGATCGGCGGACCGGTTAAAC	362
Qy	301	GGGCATCATCCCGGTGACGTATCTCCGTGACCTGGCAAAACCAAGTCGGGGGCAACGGT	360
Db	363	GGGCATCATCCCGGTGACGTATCTCCGTGACCTGGCAAAACCAAGTCGGGGGCAACGGT	422
Qy	361	ACAGGGAACGTGAATTTGGCGGAGGGAACCCCGGCC	396
Db	423	ACAGGGAACGTGACATTTGGCGGAGGGAACCCCGGCC	458

```

RESULT 4
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

	Query Match	Best Local Similarity	100.0%;	Score 396;	DB 3;	Length 4403765;
	Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ACGGCGCGGTCCGATAACTTCAGCTGTGCCAGGTGGSCAGGGATTCGCCATTCGGATC	60			
Db	151984	ACGGCGCGGTCCGATAACTTCAGCTGTGCCAGGTGGSCAGGGATTCGCCATTCGGATC	152043			
QY	61	GGGCAAGCGCATGGCGATCCGCGGGCCAGATCCGATCGGGTGGGGGGTTCACCCACCGCTTCAT	120			
Db	152044	GGGCAAGCGCATGGCGATCCGCGGGCCAGATCCGATCGGGTGGGGGGTTCACCCACCGCTTCAT	152103			

Qy	121	ATCGGGCCCTACCGCCTTCTCGCCTTGGGTGTTGTCGACAAACACGCAACGCGCGACGCA	180
Db	152104	ATTCGGGCCCTACCGCCTTCTCGCCTTGGGTGTTGTCGACAAACGCGCAACGCGCGACGCA	152163
Qy	181	GTCCAAACGGCTGGTTCGGGAGCGCTCCGGCGGCGAAGTCTCGGCATCTCCACCGGCGAGCTG	240
Db	152164	GTCCAAACGGCTGGTTCGGGAGCGCTCCGGCGGCGAAGTCTCGGCATCTCCACCGGCGAGCTG	152223
Qy	241	ATCACCGCGTGTGACGGCGCTCCGATCAACTCGGGCGACCCGCGATGGCGGACGCGCTTAAC	300
Db	152224	ATCACCGCGTGTGACGGCGCTCCGATCAACTCGGGCGACCCGCGATGGCGGACGCGCTTAAC	152283
Qy	301	GGGCATCATCCCCGGTGAAGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGACGCGT	360
Db	152284	GGGCATCATCCCCGGTGAAGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGGCGGACGCGT	152343
Qy	361	ACAGGGAAACGTGACATTGGCGCGAGGACCCCGGCC	396
Db	152344	ACAGGGAAACGTGACATTGGCGCGAGGACCCCGGCC	152379

```

RESULT 5
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	100.0%	Score 396	DB 3	Length 4411529	
Best Local Similarity	100.0%	Prod. No. 2.1e-94			
Matches 396	Conservative	0	Mismatches 0	Indels 0	Gaps 0
QY	1	ACGCCGGTCCGATAACTCCAGTGTCCACGGTGGCAGGGATTGCCATTCGGATC	60		
DB	151815	ACGCCGGGTCCGATAATTCCAGTGTCCACGGTGGCAGGGATTGCCATTCGGATC	151874		
QY	61	GGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGGTCAACCCACCGTTTCAT	120		
DB	151875	GGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGTGGGGGGTCAACCCACCGTTTCAT	151934		
QY	121	ATCGGGGCTTACCGGCTTCTCGGTTGGGTGTTGTGACAACACGGCAACGGCGCACGA	180		
DB	151935	ATCGGGGCTTACCGGCTTCTCGGTTGGGTGTTGTGACAACACGGCAACGGCGCACGA	151994		
QY	191	GTCCAAACGGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTTG	240		
DB	151995	GTCCAAACGGTGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTTG	152054		
QY	241	ATCACCGCGGTGACGGCGGCTCCGATCAACTCGGCCACCGCGATGGCGACACGCGTTTAAC	300		
DB	152055	ATCACCGCGGTGACGGCGGCTCCGATCAACTCGGCCACCGCGATGGCGACACGCGTTTAAC	152114		
QY	301	GGGCATCATCCCGGTGACGTGATCTCGGTGACCTTGGCAAAACCAAGTCGGCGGCGACGGT	360		
DB	152115	GGGCATCATCCCGGTGACGTGATCTCGGTGACCTTGGCAAAACCAAGTCGGCGGCGACGGT	152174		

QY 361 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 396  
Db 152175 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 152210

RESULT 6

US-08-818-112-4  
; Sequence 4, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 682-6031  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-112-4

Query Match 99.6%; Score 394.4; DB 3; Length 447;  
Best Local Similarity 99.7%; Pred. No. 7.6e-95;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACAGGCGGTCCGATAAATTCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 11 ACAGGCGGTCCGATAAATTCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 70  
QY 61 GGGCAGGCGATGGCGATCCGCGGCGAGATCCGATCCGGTGGGGGTCAACCAACCGTTTCAT 120  
Db 71 GGGCAGGCGATGGCGATCCGCGGCGAGATCCGATCCGGTGGGGGTCAACCAACCGTTTCAT 130  
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTCGACAAACGCGAAACGCGCGACGA 180  
Db 131 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTCGACAAACGCGAAACGCGCGACGA 190  
QY 181 GTCCACGCGTGGTGGGAGCGCTCCGCGGCGAGTCTCGGATCTCCACCGGCGACGTG 240  
Db 191 GTCCACGCGTGGTGGGAGCGCTCCGCGGCGAGTCTCGGATCTCCACCGGCGACGTG 250  
QY 241 ATCACGCGGTGCAGCGGCTCCGATCAACTCGGCCACCGGATCGCGGAGCGGCTTAAC 300

Db 251 ATCACGCGGTGCAGCGGCTCCGATCAACTCGGCCACCGGATGGCGGACCGCGCTTAAC 310  
QY 301 GGGCATCATCCGGTGAGTCACTTCGGTGACCTGGCAAAACAAAGTCGGGCGGACCGGT 360  
Db 311 GGGCATCATCCGGTGAGTCACTTCGGTGAGTGGCAAAACAAAGTCGGGCGGACCGGT 370  
QY 361 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 396  
Db 371 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 406

RESULT 7

US-08-818-111-4  
; Sequence 4, Application US/08818111  
; Patent No. 633852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,111  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-111-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
Best Local Similarity 99.7%; Pred. No. 7.6e-95;  
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACAGGCGGTCCGATAAATTCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 11 ACAGGCGGTCCGATAAATTCAGCTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 70  
QY 61 GGGCAGGCGATGGCGATCCGCGGCGAGATCCGATCCGGTGGGGGTCAACCAACCGTTTCAT 120  
Db 71 GGGCAGGCGATGGCGATCCGCGGCGAGATCCGATCCGGTGGGGGTCAACCAACCGTTTCAT 130  
QY 121 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTCGACAAACGCGAAACGCGCGACGA 180  
Db 131 ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTCGACAAACGCGAAACGCGCGACGA 190  
QY 181 GTCCACGCGTGGTGGGAGCGCTCCGCGGCGAGTCTCGGATCTCCACCGGCGACGTG 240

Db 191 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGACGCTG 250  
 Qy 241 ATCAACCGCGTCCGAGCGCTCCGATCAACTCGGCAACCGGATCGGCGGCGCTTAAC 300  
 Db 251 ATCAACCGCGTCCGAGCGCTCCGATCAACTCGGCAACCGGATCGGCGGCGCTTAAC 310  
 Qy 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTTGGCAACCAAGTCCGGCGGCGACGCGT 360  
 Db 311 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTTGGCAACCAAGTCCGGCGGCGACGCGT 370  
 Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
 Db 371 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 406

RESULT 8  
 US-09-056-556-4  
 ; Sequence 4, Application US/09056556  
 ; Patent No. 6350456  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Campos-Neto, Antonia  
 ; APPLICANT: Houghton, Raymond  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Twardzik, Daniel R.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Hendrickson, Ronald C.  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104-7092  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA: US/09/056,556  
 ; FILING DATE: 07-APR-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maki, David J.  
 ; REGISTRATION NUMBER: 31,392  
 ; REFERENCE/DOCKET NUMBER: 210121.457  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 447 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-09-056-556-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
 Best Local Similarity 99.7%; Pred. No. 7.6e-95;  
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACAGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCAATTCGGATC 60  
 Db 11 ACAGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCAATTCGGATC 70  
 Qy 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCGGCGGTTCACCCACCGGTCAT 120  
 Db 71 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCGGCGGTTCACCCACCGGTCAT 130  
 Qy 121 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTTCGACAAACACGCGCAACGCGGACGA 180  
 Db 131 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTTCGACAAACACGCGCAACGCGGACGA 190

Qy 181 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGACGCTG 240  
 Db 191 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGGACGCTG 250  
 Qy 241 ATCAACCGCGTCCGAGCGCTCCGATCAACTCGGCAACCGGATCGGCGGCGCTTAAC 300  
 Db 251 ATCAACCGCGTCCGAGCGCTCCGATCAACTCGGCAACCGGATCGGCGGCGCTTAAC 310  
 Qy 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTTGGCAACCAAGTCCGGCGGCGACGCGT 360  
 Db 311 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTTGGCAACCAAGTCCGGCGGCGACGCGT 370  
 Qy 361 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 396  
 Db 371 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 406

RESULT 9  
 US-09-072-596-4  
 ; Sequence 4, Application US/09072596  
 ; Patent No. 6458366  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Campos-Neto, Antonia  
 ; APPLICANT: Houghton, Raymond  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Twardzik, Daniel R.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Hendrickson, Ronald C.  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104-7092  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA: US/09/072,596  
 ; FILING DATE: 05-MAY-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maki, David J.  
 ; REGISTRATION NUMBER: 31,392  
 ; REFERENCE/DOCKET NUMBER: 210121.417C9  
 ; TELEPHONE: (206) 622-4900  
 ; TELEFAX: (206) 682-6031  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 447 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-09-072-596-4

Query Match 99.6%; Score 394.4; DB 4; Length 447;  
 Best Local Similarity 99.7%; Pred. No. 7.6e-95;  
 Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACAGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCAATTCGGATC 60  
 Db 11 ACAGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGCAATTCGGATC 70  
 Qy 61 GGGCAGGCGATGGCGATCGCGGCCAGATCCGATCGGCGGCGGTTCACCCACCGGTCAT 120

TREATM

TUBERCULOSIS





TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-111-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.2e-94;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCAGAGTGTCCAGGTGGGACAGGATTGCGCAATTCGATC 60  
DB 758 ACGGCGCGTCCGATAAATTCAGAGTGTCCAGGTGGGACAGGATTGCGCAATTCGATC 817  
QY 61 GGGCAGCGGATGCGATCCGCGGCCAGATCCGATCCGGTGGGGGTTCACCCACGTTTCAT 120  
DB 818 GGGCAGCGGATGCGATCCGCGGCCAAATCCGATCCGGTGGGGGTTCACCCACGTTTCAT 877  
QY 121 ATCGGGCTTACCGCTTCCTCCGCTTTGGGTGTGTCACAAACGCAACGCGCACGA 180  
DB 878 ATCGGGCTTACCGCTTCCTCCGCTTTGGGTGTGTCACAAACGCAACGCGCACGA 937  
QY 181 GTCACACGCGTGTCCGAGAGCGTCCGCGGCGCAAGTCTCGGATCTCCACCGGCGAGTG 240  
DB 938 GTCACACGCGTGTCCGAGAGCGTCCGCGGCGCAAGTCTCGGATCTCCACCGGCGAGTG 997  
QY 241 ATCACGCGGTGCACGCGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 300  
DB 998 ATCACGCGGTGCACGCGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 1057  
QY 301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGACGCGT 360  
DB 1058 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAAACCAAGTCGGGCGGACGCGT 1117  
QY 361 ACAGGGAACTGACATTTGGCGGAGGACCCCGCGCC 396  
DB 1118 ACAGGGAACTGACATTTGGCGGAGGACCCCGCGCC 1153

RESULT 13  
US-09-056-556-17  
Sequence 17, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056.556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 21.0121.457

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-056-556-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.2e-94;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGAGGTGGCGAGGATTCGCAATCCGATC 60  
DB 758 ACGGCGCGTCCGATAAATTCAGCTGTCAGAGGTGGCGAGGATTCGCAATCCGATC 817  
QY 61 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCA 120  
DB 818 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCA 877  
QY 121 ATCGGGCTTACCGCTTCTCGCTTGGGTGTTGCGACAAACGCGCAACGGCGCACGA 180  
DB 878 ATCGGGCTTACCGCTTCTCGCTTGGGTGTTGCGACAAACGCGCAACGGCGCACGA 937  
QY 181 GTCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 240  
DB 938 GTCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 997  
QY 241 ATCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 300  
DB 998 ATCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 1057  
QY 301 GGGCATCATCCCGTGACGTCTCGGTGACCTGCGTGAACCTGCGCAACCAAGTCGGCGCGACGCGT 360  
DB 1058 GGGCATCATCCCGTGACGTCTCGGTGACCTGCGTGAACCTGCGCAACCAAGTCGGCGCGACGCGT 1117

RESULT 14  
US-09-072-596-17  
Sequence 17, Application US/09072596  
Patent No. 6459365  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedwick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-056-556-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.2e-94;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGAGGTGGCGAGGATTCGCAATCCGATC 60  
DB 758 ACGGCGCGTCCGATAAATTCAGCTGTCAGAGGTGGCGAGGATTCGCAATCCGATC 817  
QY 61 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCA 120  
DB 818 GGGCAGCGATGGCGATCGCGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCA 877  
QY 121 ATCGGGCTTACCGCTTCTCGCTTGGGTGTTGCGACAAACGCGCAACGGCGCACGA 180  
DB 878 ATCGGGCTTACCGCTTCTCGCTTGGGTGTTGCGACAAACGCGCAACGGCGCACGA 937  
QY 181 GTCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 240  
DB 938 GTCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 997  
QY 241 ATCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 300  
DB 998 ATCCAGCGGTGTCGGGCGCTCCGATCAATCGGCGCACCGGATCGGCGCGACGTG 1057  
QY 301 GGGCATCATCCCGTGACGTCTCGGTGACCTGCGTGAACCTGCGCAACCAAGTCGGCGCGACGCGT 360  
DB 1058 GGGCATCATCCCGTGACGTCTCGGTGACCTGCGTGAACCTGCGCAACCAAGTCGGCGCGACGCGT 1117

RESULT 15  
US-09-072-967-17  
Sequence 17, Application US/09072967  
Patent No. 6592877  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedwick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington

COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-967-17

Query Match 98.8%; Score 391.2; DB 4; Length 1872;  
Best Local Similarity 99.2%; Pred. No. 7.2e-94;  
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ACGGCGCGTCCGATTAATCCAGCTGTCACAGGTGGGAGGATTCGCCATTCCGATC 60  
DB 758 ACGGCGCGTCCGATTAATCCAGCTGTCACAGGTGGGAGGATTCGCCATTCCGATC 817  
QY 61 GGGCAGCGATGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCCACCGTTAT 120  
DB 818 GGGCAGCGATGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCCACCGTTAT 877  
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGTGGTTCGACAAACGCGACGCGACGA 180  
DB 878 ATCGGCGCTACCGCTTCCTCGGCTGGGTGGTTCGACAAACGCGACGCGACGA 937  
QY 181 GTCCAAACGCGTGGTTCGGAGCGCTCCGGCGGCGAAGTCTCGGCATCTCCACCGCGACGTG 240  
DB 938 GTCCAAACGCGTGGTTCGGAGCGCTCCGGCGGCGAAGTCTCGGCATCTCCACCGCGACGTG 997  
QY 241 ATACCGCGGTGACAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 300  
DB 998 ATACCGCGGTGACAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 1057  
QY 301 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCGT 360  
DB 1058 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCGT 1117  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
DB 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

RESULT 16  
US-09-636-215-822  
; Sequence 822, Application US/09636215  
; Patent No. 6620922  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.42717C17  
; CURRENT APPLICATION NUMBER: US/09/636,215  
; CURRENT FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 852  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-636-215-822

Query Match 89.4%; Score 354; DB 4; Length 675;  
Best Local Similarity 95.7%; Pred. No. 3.4e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
QY 1 ACGGCGCGTCCGATTAATCCAGCTGTCACAGGTGGGAGGATTCGCCATTCCGATC 60  
DB 22 ACGGCGCGTCCGATTAATCCAGCTGTCACAGGTGGGAGGATTCGCCATTCCGATC 81  
QY 61 GGGCAGCGATGCGGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCCACCGTTAT 120  
DB 82 GGGCAGCGATGCGGATCGCGGCGCAGATC-----AAGCTTCCACCGTTAT 129  
QY 121 ATCGGCGCTACCGCTTCCTCGGCTGGGTGGTTCGACAAACGCGACGCGACGA 180  
DB 130 ATCGGCGCTACCGCTTCCTCGGCTGGGTGGTTCGACAAACGCGACGCGACGA 189  
QY 181 GTCCAAACGCGTGGTTCGGAGCGCTCCGGCGGCGAAGTCTCGGCATCTCCACCGCGACGTG 240  
DB 190 GTCCAAACGCGTGGTTCGGAGCGCTCCGGCGGCGAAGTCTCGGCATCTCCACCGCGACGTG 249  
QY 241 ATACCGCGGTGACAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 300  
DB 250 ATACCGCGGTGACAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 309  
QY 301 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCGT 360  
DB 310 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGCGCGACGCGT 369  
QY 361 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 396  
DB 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 405

RESULT 17  
US-09-685-166A-822  
; Sequence 822, Application US/09685166A  
; Patent No. 6630305  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel

```

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: DIAGNOSIS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-822

Query Match      89.4%; Score 354; DB 4; Length 675;
Best Local Similarity 95.7%; Pred. No. 3.4e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGGAGTGGGAGGATTCGCAATTCGGATC 60
DB 22 ACGGCGCGTCCGATAAATTCAGCTGTCAGGAGTGGGAGGATTCGCAATTCGGATC 81
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCGGATCGGATCGGATCGGATCGGAT 120
DB 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCGGATCGGATCGGATCGGATCGGAT 129
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGCACGA 180
DB 130 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGCACGA 189
QY 181 GTCCAAACGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 240
DB 190 GTCCAAACGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 249
QY 241 ATCACCGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 300
DB 250 ATCACCGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 309
QY 301 GGGCATCATCCCGTGGAGTCTCGGATCACTCGGATCACTCGGATCACTCGGAT 360
DB 310 GGGCATCATCCCGTGGAGTCTCGGATCACTCGGATCACTCGGATCACTCGGAT 369
QY 361 ACAGGGAACGTGACATTCGCGGAGGAGCGGAGGAGCGGAGCGGAGCGGAG 396
DB 370 ACAGGGAACGTGACATTCGCGGAGGAGCGGAGGAGCGGAGCGGAGCGGAG 405

RESULT 19
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-353

Query Match      89.4%; Score 354; DB 4; Length 900;
Best Local Similarity 95.7%; Pred. No. 3.6e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGGAGTGGGAGGATTCGCAATTCGGATC 60
DB 22 ACGGCGCGTCCGATAAATTCAGCTGTCAGGAGTGGGAGGATTCGCAATTCGGATC 81

```

```

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: DIAGNOSIS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-822

Query Match      89.4%; Score 354; DB 4; Length 675;
Best Local Similarity 95.7%; Pred. No. 3.4e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGGAGTGGGAGGATTCGCAATTCGGATC 60
DB 22 ACGGCGCGTCCGATAAATTCAGCTGTCAGGAGTGGGAGGATTCGCAATTCGGATC 81
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCGGATCGGATCGGATCGGATCGGAT 120
DB 82 GGGCAGGCGATGGCGATCGCGGCGCAGATCGGATCGGATCGGATCGGATCGGAT 129
QY 121 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGCACGA 180
DB 130 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGCACGA 189
QY 181 GTCCAAACGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 240
DB 190 GTCCAAACGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 249
QY 241 ATCACCGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 300
DB 250 ATCACCGCGTGGTGGGAGCGCTCGGAGCGCTCGGAGCGCTCGGAGCGCTG 309
QY 301 GGGCATCATCCCGTGGAGTCTCGGATCACTCGGATCACTCGGATCACTCGGAT 360
DB 310 GGGCATCATCCCGTGGAGTCTCGGATCACTCGGATCACTCGGATCACTCGGAT 369
QY 361 ACAGGGAACGTGACATTCGCGGAGGAGCGGAGGAGCGGAGCGGAGCGGAG 396
DB 370 ACAGGGAACGTGACATTCGCGGAGGAGCGGAGGAGCGGAGCGGAGCGGAG 405

RESULT 18
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822

```

QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGCGGCGCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCTACCGGCTTCCTCGGCTTGGGTGTTTGTGCAACAACGCGCAACGCGGCGACGA 180  
Db 130 ATCGGGCTACCGGCTTCCTCGGCTTGGGTGTTTGTGCAACAACGCGCAACGCGGCGACGA 189  
QY 181 GTCCAAACGCGTGGTGGGAGGCTCCGCGCGCAAGTCTCGCATCTCCGCACTCTCCGCGGCGACGTG 240  
Db 190 GTCCAAACGCGTGGTGGGAGGCTCCGCGCGCAAGTCTCGCATCTCCGCGGCGACGTG 249  
QY 241 ATCACCGCGTTCGACGCGCTCCGATCAATTCGCGCCACCGGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCACCGCGTTCGACGCGCTCCGATCAATTCGCGCCACCGGATGGCGGACGCGCTTAAAC 309  
QY 301 GGGCATCATCCGCTGAGTCACTCTCGTGACCTGGCAACCAAGTCCGAGTGGGCGGCGACGCGT 360  
Db 310 GGGCATCATCCGCTGAGTCACTCTCGTGACCTGGCAACCAAGTCCGAGTGGGCGGCGACGCGT 369  
QY 361 ACAGGAACGTGACATTTGGCCGAGGACCCCGCGCC 396  
Db 370 ACAGGAACGTGACATTTGGCCGAGGACCCCGCGCC 405

## RESULT 20

US-09-606-421B-353  
; Sequence 353, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-606-421B-353

Query Match 89.4%; Score 354; DB 4; Length 900;  
Best Local Similarity 95.7%; Pred. No. 3.6e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
QY 1 ACGGCGGCTCGGATTAATTCAGCTGTCACAGGTTGGGCGAGGATTCGCCATTCGATC 60  
Db 22 ACGGCGGCTCGGATTAATTCAGCTGTCACAGGTTGGGCGAGGATTCGCCATTCGATC 81  
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGCGGCGCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCTTACCGCTTCCTCGGCTTGGGTGTTGTGCAACAACGCGCAACGCGGCGACGA 180  
Db 130 ATCGGGCTTACCGCTTCCTCGGCTTGGGTGTTGTGCAACAACGCGCAACGCGGCGACGA 189  
QY 181 GTCCAAACGCGTGGTGGGAGGCTCCGCGCGCAAGTCTCGCATCTCCGCGGCGACGTG 240  
Db 190 GTCCAAACGCGTGGTGGGAGGCTCCGCGCGCAAGTCTCGCATCTCCACCGGCGACGTG 249

QY 241 ATCACCGGCTCGACGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCACCGGCTCGACGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAAC 309  
QY 301 GGGCATCATCCGCTGAGTCACTCTCGTGACCTGGCAACCAAGTCCGAGTGGGCGGCGACGCGT 360  
Db 310 GGGCATCATCCGCTGAGTCACTCTCGTGACCTGGCAACCAAGTCCGAGTGGGCGGCGACGCGT 369  
QY 361 ACAGGAACGTGACATTTGGCCGAGGACCCCGCGCC 396  
Db 370 ACAGGAACGTGACATTTGGCCGAGGACCCCGCGCC 405

## RESULT 21

US-09-636-215-834  
; Sequence 834, Application US/09636215  
; Patent No. 6620922  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.42717C17  
; CURRENT APPLICATION NUMBER: US/09/636,215  
; CURRENT FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 852  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 834  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-636-215-834

Query Match 89.4%; Score 354; DB 4; Length 915;  
Best Local Similarity 95.7%; Pred. No. 3.6e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGGCTCGGATTAATTCAGCTGTCACAGGTTGGGCGAGGATTCGCCATTCGATC 60  
Db 22 ACGGCGGCTCGGATTAATTCAGCTGTCACAGGTTGGGCGAGGATTCGCCATTCGATC 81  
QY 61 GGGCAGGCGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCAACCGCGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGCGGCGCAGATC-----AAGCTTCCACCGTTTCAT 129  
QY 121 ATCGGGCTTACCGCTTCCTCGGCTTGGGTGTTGTGCAACAACGCGCAACGCGGCGACGA 180  
Db 130 ATCGGGCTTACCGCTTCCTCGGCTTGGGTGTTGTGCAACAACGCGCAACGCGGCGACGA 189  
QY 181 GTCCAAACGCGTGGTGGGAGGCTCCGCGGCGCAAGTCTCGCATCTCCGCGGCGACGTG 240  
Db 190 GTCCAAACGCGTGGTGGGAGGCTCCGCGGCGCAAGTCTCGCATCTCCACCGGCGACGTG 249  
QY 241 ATCACCGCGTTCGACGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAAC 300  
Db 250 ATCACCGCGTTCGACGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAAC 309  
QY 301 GGGCATCATCCCGTGAACGTTCATCTCGGTGAACCTGGCAACCAAGTCCGAGTGGGCGGCGACGCGT 360

Db 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTTGGCAACCAAGTCGGCGGCACGCGT 369  
QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 405

RESULT 22  
US-09-685-166A-834  
; Sequence 834, Application US/09685166A  
; Patent No. 6630305  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Harlock, Jennifer L.  
; APPLICANT: Harlock, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C21  
; CURRENT APPLICATION NUMBER: US/09/685,166A  
; CURRENT FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 898  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 834  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-685-166A-834

Query Match 89.4%; Score 354; DB 4; Length 915;  
Best Local Similarity 95.7%; Pred. No. 3.6e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGGTTGGGAGGATTCGCCATTCGGATC 60  
Db 22 ACGGCGCGTCCGATAAATTCAGCTGTCAGGTTGGGAGGATTCGCCATTCGGATC 81

QY 61 GGGCAGGCGATGCGGATCCGCGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCA 120  
Db 82 GGGCAGGCGATGCGGATCCGCGGCCAGATC-----AAGCTTCCACCGTTTCA 129

QY 121 ATCGGCGCTACCGCTTCTCGGTGGTGTGTCGACAAACGCGCAACGCGCACGA 180  
Db 130 ATCGGCGCTACCGCTTCTCGGTGGTGTGTCGACAAACGCGCAACGCGCACGA 189

QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGAGCGTCCGCGGAGCGTCCGCGGAGCGT 240  
Db 190 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGAGCGTCCGCGGAGCGTCCGCGGAGCGT 249

QY 241 ATCAACGCGGTGCGAGCGCGTCCGATCAACTCGGCGCACCGCGATGCGGCGGCTTAAC 300  
Db 250 ATCAACGCGGTGCGAGCGCGTCCGATCAACTCGGCGCACCGCGATGCGGCGGCTTAAC 309

QY 301 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTTCGGCGGCGCGGT 360  
Db 310 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTTCGGCGGCGCGGT 369

QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396

Db 370 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 405

RESULT 23  
US-09-736-457-1861  
; Sequence 1861, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Panger, Gary  
; APPLICANT: Vedwick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1861  
; LENGTH: 945  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1861

Query Match 89.4%; Score 354; DB 4; Length 945;  
Best Local Similarity 95.7%; Pred. No. 3.7e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACGGCGCGTCCGATAAATTCAGCTGTCAGGTTGGGAGGATTCGCCATTCGGATC 60  
Db 22 ACGGCGCGTCCGATAAATTCAGCTGTCAGGTTGGGAGGATTCGCCATTCGGATC 81

QY 61 GGGCAGGCGATGCGGATCCGCGGCCAGATCCGATCGGTTGGGGGTCAACCCACCGTTTCA 120  
Db 82 GGGCAGGCGATGCGGATCCGCGGCCAGATC-----AAGCTTCCACCGTTTCA 129

QY 121 ATCGGCGCTACCGCTTCTCGGTGGTGTGTCGACAAACGCGCAACGCGCACGA 180  
Db 130 ATCGGCGCTACCGCTTCTCGGTGGTGTGTCGACAAACGCGCAACGCGCACGA 189

QY 181 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGAGCGTCCGCGGAGCGTCCGCGGAGCGT 240  
Db 190 GTCCAAACGCGTGTGCGGAGCGCTCCGCGGAGCGTCCGCGGAGCGTCCGCGGAGCGT 249

QY 241 ATCAACGCGGTGCGAGCGCGTCCGATCAACTCGGCGCACCGCGATGCGGCGGCTTAAC 300  
Db 250 ATCAACGCGGTGCGAGCGCGTCCGATCAACTCGGCGCACCGCGATGCGGCGGCTTAAC 309

QY 301 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTTCGGCGGCGCGGT 360  
Db 310 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTTCGGCGGCGCGGT 369

QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 405

RESULT 24  
US-09-643-597-351  
; Sequence 351, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.

```
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-351

Query Match      89.4%; Score 354; DB 4; Length 1012;
Best Local Similarity 95.7%; Pred. No. 3.7e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACAGCCGGCTCCGATAAATTCAGCTGTCCAGAGGTGGGAGGATTCGGCATTCGGATC 60
DB 67 ACAGCCGGCTCCGATAAATTCAGCTGTCCAGAGGTGGGAGGATTCGGCATTCGGATC 126
QY 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTGACCCACCGTTTCA 120
DB 127 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGGCGAGATC-----AAGCTTCCACCGTTTCA 174
QY 121 ATCCGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGGCGACGA 180
DB 175 ATCCGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGGCGACGA 234
QY 181 GTCCAAACGCGTGGTGGAGCGCTCCGGGCGAGATCCGATCGGGTGGGGGTGACCCACCGTTTCA 240
DB 235 GTCCAAACGCGTGGTGGAGCGCTCCGGGCGAGATCCGATCGGGTGGGGGTGACCCACCGTTTCA 294
QY 241 ATCCAGCGGTGCGAGCGGCTCCGATCAACTCGGCCACCGCAACGCGGCGGCTTAAC 300
DB 295 ATCCAGCGGTGCGAGCGGCTCCGATCAACTCGGCCACCGCAACGCGGCGGCTTAAC 354
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 360
DB 355 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 414
QY 361 ACAGGAAACGTGACATTTGGCGGGGACCCCGGCC 396
DB 415 ACAGGAAACGTGACATTTGGCGGGGACCCCGGCC 450

RESULT 25
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 651315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-351

Query Match      89.4%; Score 354; DB 4; Length 1012;
Best Local Similarity 95.7%; Pred. No. 3.7e-84;
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

QY 1 ACAGCCGGCTCCGATAAATTCAGCTGTCCAGAGGTGGGAGGATTCGGCATTCGGATC 60
DB 67 ACAGCCGGCTCCGATAAATTCAGCTGTCCAGAGGTGGGAGGATTCGGCATTCGGATC 126
QY 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTGACCCACCGTTTCA 120
DB 127 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGGCGAGATC-----AAGCTTCCACCGTTTCA 174
QY 121 ATCCGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGGCGACGA 180
DB 175 ATCCGGCTACCGCTTCTCGGCTTGGGTGTGTCGACAAACGCAACGCGGCGACGA 234
QY 181 GTCCAAACGCGTGGTGGAGCGCTCCGGGCGAGATCCGATCGGGTGGGGGTGACCCACCGTTTCA 240
DB 235 GTCCAAACGCGTGGTGGAGCGCTCCGGGCGAGATCCGATCGGGTGGGGGTGACCCACCGTTTCA 294
QY 241 ATCCAGCGGTGCGAGCGGCTCCGATCAACTCGGCCACCGCAACGCGGCGGCTTAAC 300
DB 295 ATCCAGCGGTGCGAGCGGCTCCGATCAACTCGGCCACCGCAACGCGGCGGCTTAAC 354
QY 301 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 360
DB 355 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGACGCGT 414
QY 361 ACAGGAAACGTGACATTTGGCGGGGACCCCGGCC 396
DB 415 ACAGGAAACGTGACATTTGGCGGGGACCCCGGCC 450

RESULT 26
US-09-636-215-851
; Sequence 851, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-636-215-851

Query Match 89.4%; Score 354; DB 4; Length 1203;  
Best Local Similarity 95.7%; Pred. No. 3.8e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
  
Qy 1 ACGGCGCGTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 22 ACGGCGCGTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 81  
  
Qy 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTCAACCCAGCTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTCAACCCAGCTTCAT 129  
  
Qy 121 ATCGGGCTTACCGCTTCTCGGTGGGTGTGCGACAAACAAACGCAACCGCGCACGA 180  
Db 130 ATCGGGCTTACCGCTTCTCGGTGGGTGTGCGACAAACAAACGCGCACGA 189  
  
Qy 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGAGATCTCGGCATCTCCACCGGCGAGTG 240  
Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGAGATCTCGGCATCTCCACCGGCGAGTG 249  
  
Qy 241 ATCAGCGGTCGAGCGGCTCCGATCACTCGGCACCGGATGGGCGCGCTTAAC 300  
Db 250 ATCAGCGGTCGAGCGGCTCCGATCACTCGGCACCGGATGGGCGCGCTTAAC 309  
  
Qy 301 GGGCATCATCCCGGTGACGTCTCGGTGACCTCGGCAACCAAGTCGGGCGGCGAGTG 360  
Db 310 GGGCATCATCCCGGTGACGTCTCGGTGACCTCGGCAACCAAGTCGGGCGGCGAGTG 369  
  
Qy 361 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 405

RESULT 27

US-09-685-166A-851  
; Sequence 851, Application US/09685166A  
; Patent No. 6630305  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C21  
; CURRENT APPLICATION NUMBER: US/09/685,166A  
; CURRENT FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 898  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 851  
; LENGTH: 1203  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-685-166A-851  
  
Query Match 89.4%; Score 354; DB 4; Length 1203;  
Best Local Similarity 95.7%; Pred. No. 3.8e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCGCGTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 22 ACGGCGCGTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 81  
  
Qy 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTCAACCCAGCTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTCAACCCAGCTTCAT 129  
  
Qy 121 ATCGGGCTTACCGCTTCTCGGTGGGTGTGCGACAAACAAACGCAACCGCGCACGA 180  
Db 130 ATCGGGCTTACCGCTTCTCGGTGGGTGTGCGACAAACAAACGCGCACGA 189  
  
Qy 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGAGATCTCGGCATCTCCACCGGCGAGTG 240  
Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGAGATCTCGGCATCTCCACCGGCGAGTG 249  
  
Qy 241 ATCAGCGGTCGAGCGGCTCCGATCACTCGGCACCGGATGGGCGCGCTTAAC 300  
Db 250 ATCAGCGGTCGAGCGGCTCCGATCACTCGGCACCGGATGGGCGCGCTTAAC 309  
  
Qy 301 GGGCATCATCCCGGTGACGTCTCGGTGACCTCGGCAACCAAGTCGGGCGGCGAGTG 360  
Db 310 GGGCATCATCCCGGTGACGTCTCGGTGACCTCGGCAACCAAGTCGGGCGGCGAGTG 369  
  
Qy 361 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGAACTGACATTTGGCGGAGGACCCCGGCC 405

RESULT 28

US-09-620-412C-348  
; Sequence 348, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: Fast-Seq for Windows Version 3.0/4.0  
; SEQ ID NO 348  
; LENGTH: 1464  
; TYPE: DNA  
; ORGANISM: Chlamydia trachomatis  
US-09-620-412C-348

Query Match 89.4%; Score 354; DB 4; Length 1464;  
Best Local Similarity 95.7%; Pred. No. 4e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;

Qy 1 ACGGCGCGTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 60  
Db 22 ACGGCGCGTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGCCATTCCGATC 81  
  
Qy 61 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTCAACCCAGCTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGGGGCGAGATCCGATCGGGTGGGGGTCAACCCAGCTTCAT 129  
  
Qy 121 ATCGGGCTTACCGCTTCTCGGTGGGTGTGCGACAAACAAACGCAACCGCGCACGA 180  
Db 130 ATCGGGCTTACCGCTTCTCGGTGGGTGTGCGACAAACAAACGCGCACGA 189  
  
Qy 181 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGAGATCTCGGCATCTCCACCGGCGAGTG 240  
Db 190 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGAGATCTCGGCATCTCCACCGGCGAGTG 249  
  
Qy 241 ATCAGCGGTCGAGCGGCTCCGATCACTCGGCACCGGATGGGCGGCGCTTAAC 300  
Db 250 ATCAGCGGTCGAGCGGCTCCGATCACTCGGCACCGGATGGGCGGCGCTTAAC 309



QY 301 GGGCATCATCCGGTACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGCACGGCT 360  
Db 310 GGGCATCATCCGGTACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGCACGGCT 369  
QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 405

RESULT 29

US-09-598-419-348  
; Sequence 348, Application US/09598419  
; Patent No. 656956  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Scholler, John  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C6  
; CURRENT APPLICATION NUMBER: US/09/598,419  
; CURRENT FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 357  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 348  
; LENGTH: 1464  
; TYPE: DNA  
; ORGANISM: Chlamydia trachomatis  
US-09-598-419-348

Query Match 89.4%; Score 354; DB 4; Length 1464;  
Best Local Similarity 95.7%; Pred. No. 4e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
QY 1 ACGGCGCGTCCGATACTTCAGCTGTCAGAGGTGGGAGGATTCGGCATTCGGATC 60  
Db 22 ACGGCGCGTCCGATACTTCAGCTGTCAGAGGTGGGAGGATTCGGCATTCGGATC 81  
QY 61 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTCGGGGGTCAACCGGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTCGGGGGTCAACCGGTTTCAT 129  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGACGA 180  
Db 130 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGACGA 189  
QY 181 GTCCAAACGCGTGGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 190 GTCCAAACGCGTGGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249  
QY 241 ATCACCGGCGTCCGATCAACTCGGCGCACCGGATGGCGACCGGCTTTAAC 300  
Db 250 ATCACCGGCGTCCGATCAACTCGGCGCACCGGATGGCGACCGGCTTTAAC 309  
QY 301 GGGCATCATCCGGTACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGCT 360  
Db 310 GGGCATCATCCGGTACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGCT 369  
QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 405

RESULT 30

US-09-620-412C-332  
; Sequence 332, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 332  
; LENGTH: 1557  
; TYPE: DNA  
; ORGANISM: Chlamydia trachomatis  
US-09-620-412C-332

Query Match 89.4%; Score 354; DB 4; Length 1557;  
Best Local Similarity 95.7%; Pred. No. 4.1e-84;  
Matches 379; Conservative 0; Mismatches 5; Indels 12; Gaps 1;  
QY 1 ACGGCGCGTCCGATACTTCAGCTGTCAGAGGTGGGAGGATTCGGCATTCGGATC 60  
Db 22 ACGGCGCGTCCGATACTTCAGCTGTCAGAGGTGGGAGGATTCGGCATTCGGATC 81  
QY 61 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTCGGGGGTCAACCGGTTTCAT 120  
Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTCGGGGGTCAACCGGTTTCAT 129  
QY 121 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGACGA 180  
Db 130 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACAAACGCAACGGCGACGA 189  
QY 181 GTCCAAACGCGTGGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240  
Db 190 GTCCAAACGCGTGGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249  
QY 241 ATCACCGGCGTCCGATCAACTCGGCGCACCGGATGGCGACCGGCTTTAAC 300  
Db 250 ATCACCGGCGTCCGATCAACTCGGCGCACCGGATGGCGACCGGCTTTAAC 309  
QY 301 GGGCATCATCCGGTACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGCT 360  
Db 310 GGGCATCATCCGGTACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGCT 369  
QY 361 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 396  
Db 370 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 405

Search completed: April 30, 2004, 07:46:54  
Job time : 72 secs

29: gb\_ges2:\*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 22:25:27 ; Search time 2585.37 Seconds  
(without alignments)  
1524.660 Million cell updates/sec

Title: US-09-684-215B-4  
Perfect score: 675  
Sequence: 1 TAAADNQLSQGGGFAPI.....QTKSGGTRGNTVLAEGPPA 132

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spoos/US09684215/runat\_29042004\_061305\_13200/app\_query.fasta\_1.1180  
-DB=EST -QFMT=fastap -SUFFIX=arst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09684215 @CGN 1 1 5167 @runat\_29042004\_061305\_13200 -NCPU=3  
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_ges1:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query %	Length	DB	ID	Description
1	145.5	21.6	289	29	U82114	U82114	U82114 U82114 orde
2	126.5	18.7	1438	28	BH770798	BH770798	BH770798 LLMT-a254
3	120.5	18.1	959	28	BZ549048	BZ549048	BZ549048 pacsl-80
4	120.5	17.9	726	14	CB679186	CB679186	CB679186 OSUNEF02F
5	117.5	17.4	590	14	CB925602	CB925602	CB925602 ABAL-21_F
6	115.5	17.1	807	14	CA228900	CA228900	CA228900 SCQFL305
7	114	16.9	603	28	AZ934428	AZ934428	AZ934428 BJ_Ba000
8	113	16.7	758	13	B0514888	B0514888	B0514888 EST622303
9	108	16.0	982	14	CD080132	CD080132	CD080132 MA3-9999U
10	106.5	15.8	1033	28	BZ561390	BZ561390	BZ561390 pacs2-164
11	103.5	15.3	719	28	AZ933900	AZ933900	AZ933900 BJ_Ba000
12	103.5	15.3	947	28	BZ549047	BZ549047	BZ549047 pacsl-80
13	103	15.3	726	28	AQ989479	AQ989479	AQ989479 rfc00035
14	102	15.1	765	9	AJ558965	AJ558965	AJ558965 AJ558965
15	100.5	14.9	859	14	CA480614	CA480614	CA480614 AGENCOURT
16	99	14.7	423	9	AI597611	AI597611	AI597611 tnl5f02.x
17	97.5	14.4	499	12	BI350520	BI350520	BI350520 fr32b05.Y
18	97.5	14.4	870	14	CA473840	CA473840	CA473840 AGENCOURT
19	97	14.4	444	13	CA148171	CA148171	CA148171 SCEERZ101
20	97	14.4	543	12	BI721127	BI721127	BI721127 1031054B1
21	96	14.2	546	10	AW285510	AW285510	AW285510 LG1_241.E
22	96	14.2	551	10	AW285527	AW285527	AW285527 LG1_241.G
23	96	14.2	1093	13	BU557763	BU557763	BU557763 AGENCOURT
24	96	14.2	1394	28	BZ576076	BZ576076	BZ576076 mesh2_4761
25	95.5	14.1	600	14	CF787130	CF787130	CF787130 854399.MA
26	95.5	14.1	647	13	BX674871	BX674871	BX674871 BX674871
27	95.5	14.1	677	12	BI378928	BI378928	BI378928 BFLG1.000
28	95	14.1	881	14	CK161112	CK161112	CK161112 FGAS04279
29	95	14.1	909	14	CA975497	CA975497	CA975497 AGENCOURT
30	94.5	14.0	728	12	BJ285591	BJ285591	BJ285591 BJ285591
31	94.5	14.0	1029	12	BI457360	BI457360	BI457360 603185681
32	94	13.9	434	28	BH638292	BH638292	BH638292 1008021F0
33	94	13.9	672	14	CF844326	CF844326	CF844326 psHB02XJ
34	94	13.9	1321	13	BU509638	BU509638	BU509638 AGENCOURT
35	93.5	13.9	600	28	CC345457	CC345457	CC345457 OGQAF70TH
36	93.5	13.9	766	14	CF997044	CF997044	CF997044 AGENCOURT
37	93.5	13.9	907	14	CA471201	CA471201	CA471201 AGENCOURT
38	93.5	13.9	913	14	CF268501	CF268501	CF268501 AGENCOURT
39	93	13.8	689	12	BM448307	BM448307	BM448307 DSA024D06
40	93	13.8	951	14	CA474469	CA474469	CA474469 AGENCOURT
41	92.5	13.7	452	12	BM874057	BM874057	BM874057 laa07a02.
42	92.5	13.7	709	29	CG985905	CG985905	CG985905 pastbac00
43	92.5	13.7	860	14	CA471291	CA471291	CA471291 AGENCOURT
44	92.5	13.7	949	14	CF243195	CF243195	CF243195 AGENCOURT
45	92.5	13.7	1117	14	CK208964	CK208964	CK208964 FGAS02069

# ALIGNMENTS

RESULT 1  
U82114  
LOCUS U82114  
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone  
ACCESSION U82114  
VERSION U82114.1  
KEYWORDS GI:3647212  
SOURCE Mycobacterium leprae  
ORGANISM Mycobacterium leprae  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
REFERENCE 1 (bases 1 to 289)

**AUTHORS** Silbaq,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.  
**TITLE** Characterization of a 34-kilodalton protein of *Mycobacterium leprae* that is isologous to the immunodominant 34-kilodalton antigen of *Mycobacterium paratuberculosis*  
**JOURNAL** Infect. Immun. 66 (11), 5576-5579 (1998)  
**MEDLINE** 99003183  
**PUBMED** 9784577  
**COMMENT** Contact: Silbaq FS  
Microbiology  
Colorado State University  
Fort Collins, CO 80523, USA  
Eiglmeyer,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae*. Mol. Microbiol. 7 (2), 197-206 (1993)  
Class: unknown.

**FEATURES**  
source  
1. -289  
/organism="Mycobacterium leprae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1769"  
/clone="cosmid L-373; contig 64"  
/clone\_lib="ordered cosmid library"

**ORIGIN**  
Alignment Scores:  
Pred. No.: 8,758-05 Length: 289  
Score: 145.50 Matches: 35  
Percent Similarity: 58.82% Conservative: 15  
Best Local Similarity: 41.18% Mismatches: 34  
Query Match: 21.56% Indels: 1  
DB: 29 Gaps: 1

US-09-684-215B-4 (1-132) x U82114 (1-289)  
QY 45 AlaPheLeuGlyValValAspAsnAenglyAsn---GlyAlaArgValGlnArg 63  
Db 7 GCGTGGCTGGTGTACAGTGGCCACCAAGGACCCCGCGCGCAAGATTATGGAC 66  
QY 64 ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 83  
Db 67 GTCGTGGCGGTGGTGGCGCCGCGAATGCGCGGTTCACAGGGGTGTTCTACTAAG 126  
QY 84 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 103  
Db 127 GTCACGACCGCTGATCAGTAGCGCGCGAGTGGTGGTGGTGGTGGTGGTGGTGGT 186  
QY 104 ProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsn 123  
Db 187 CCGGGTGACAAAGTGGCTGACCTATCAGATCAGATCAGTCTGGTAGCGTGGTGGTGGT 246  
QY 124 ValThrLeuAlaGlu 128  
Db 247 GTCACACTCGGCAAG 261

**RESULT 2**  
BH770798/c  
LOCUS BH770798 1438 bp DNA linear GSS 01-MAY-2002  
DEFINITION LMGtag541 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.  
ACCESSION BH770798  
VERSION BH770798.1 GI:20373755  
KEYWORDS GSS.  
SOURCE Lactococcus lactis subsp. cremoris  
ORGANISM Lactococcus lactis subsp. cremoris  
REFERENCE 1 (bases 1 to 1438)  
AUTHORS Bolerin,A., Ehrlich,S.D. and Sorokin,A.  
TITLE Studies of genomes of dairy bacteria Lactococcus lactis  
JOURNAL Sci. Aliments (2002) In press  
COMMENT Contact: Sorokin A  
Genetique Microbienne  
INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France  
Tel: 33 1 34 65 25 16  
Fax: 33 1 34 65 25 21  
Email: sorokine@jouy.inra.fr  
best homologue in strain ILL403 is hraA (95%)  
Class: shotgun  
High quality sequence start: 30  
High quality sequence stop: 1408.  
prepared by partial AluI digestion or by sonication.

**FEATURES**  
source  
1. -1438  
/organism="Lactococcus lactis subsp. cremoris"  
/mol\_type="genomic DNA"  
/strain="MG1363"  
/sub\_species="cremoris"  
/db\_xref="taxon:1359"  
/clone\_lib="MG1363 Random Sequence Tag Library"  
/notes="Vector: pSGMU2; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."

**ORIGIN**  
Alignment Scores:  
Pred. No.: 0.0656 Length: 1438  
Score: 126.50 Matches: 40  
Percent Similarity: 45.45% Conservative: 20  
Best Local Similarity: 30.30% Mismatches: 49  
Query Match: 18.74% Indels: 23  
DB: 28 Gaps: 5

US-09-684-215B-4 (1-132) x BH770798 (1-1438)  
QY 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32  
Db 479 GGTTAGGTTTCTATCCATCTATGATGATGGTAAACATCATTAATAAAGTGAAC 420  
QY 33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal 52  
Db 419 GATGGT-----AAGATTTCACGCTCTGCC---TTAGGTATTGGTATGGTT 378  
QY 53 Asp-----AsnAsnGlyAsn----- 57  
Db 377 GACATCTCAATATCAACAATGATGTTCTCAACTGAATACCTAGCAGCGTAAC 318  
QY 58 ---GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSer 76  
Db 317 GGTGGAGTGGTGTCTACTCTGTTCAAGGGGGTCTTCTGCTGCCACAGTGGTCTGAAA 258  
QY 77 ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAla 96  
Db 257 GCTGGCGATGTGATAACGAAGGTGGAGATACCGCGGTACTTCTATCAACAGACTTACA 198  
QY 97 AspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSer 116  
Db 197 AGTGCTCTTTACTCACACAATATTATGATGATGTAAGTCACTACTACTAC---CGTGAT 141  
QY 117 GlyGlyThrArgThrGlyAsnValThrLeuAlaGlu 128  
Db 140 GGTAAATCAGCCACAGCAAAATGTCAAACTCTCTAAA 105

**RESULT 3**  
BZ549048/c  
LOCUS BZ549048 959 bp DNA linear GSS 17-DEC-2002  
DEFINITION pacsl-60\_1644.s2 pacsl-60 Pseudomonas aeruginosa genomic clone  
ACCESSION BZ549048  
VERSION BZ549048.1 GI:27152629  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonas aeruginosa  
REFERENCE 1 (bases 1 to 959)  
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns, J.L., Kaul, R. and Olsen, M.V.  
 Whole-Genome-Sequence variation among multiple isolates of  
 Pseudomonas aeruginosa library  
 J. Bacteriol. (2002) In press  
 Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.  
 Location/Qualifiers  
 1. 959  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="1-60"  
 /db\_xref="taxon:287"  
 /clone="pacsi-60.1644"  
 /clone\_lib="pacsi-60"  
 /note="clinical isolate 1-60 Whole genomic shotgun  
 library."  
 FEATURES  
 source  
 1. 959  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="1-60"  
 /db\_xref="taxon:287"  
 /clone="pacsi-60.1644"  
 /clone\_lib="pacsi-60"  
 /note="clinical isolate 1-60 Whole genomic shotgun  
 library."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.101 Length: 959  
 Score: 122.00 Matches: 40  
 Percent Similarity: 47.29% Conservative: 21  
 Best Local Similarity: 31.01% Mismatches: 44  
 Query Match: 18.07% Indels: 24  
 DB: 28 Gaps: 5  
 US-09-684-215B-4 (1-132) x B2549048 (1-959)  
 QY 4 SerAspAsnPhgLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21  
 Db 378 TCGCAGATCTTCACCCGTTCCGGCGCTTCATGGCGCTTCCTTCGCCATTCGATCGAT 319  
 QY 22 GlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle 41  
 Db 318 GTCCGCGCTGACGTCGCCACGATGGAAGAGCCGCGC-----AAGGTAGTCGC 268  
 QY 42 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn----- 55  
 Db 267 GGC-----TGGCTGGCGGTGGTGTGATCCAGGAAGTGAACAGGATCTCGCGAGTCC 217  
 QY 56 -----GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70  
 Db 216 TACGGCTCGCAAGCCGTCGGCGCGTGTGGCGCACTGTGGAAGAGCGCCGCGC 157  
 QY 71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyValaProIleAsn 90  
 Db 156 GCCAGGGTGGCTGCAGGTGGCGATGTGATCATAGCTTGAACGCCAGTCTGATCAAC 97  
 QY 91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110  
 Db 96 GAGTCCGCGACCTCGCCGACCTGTGGGCAACATGAAGCCGCGGCAAGATCAACCTG 37  
 QY 111 ThrTrpGlnThrLysSerGlyGlyThr 119  
 Db 36 -----CACGGGGGATCC 25  
 RESULT 4  
 CB679186/c 726 bp mRNA linear EST 09-APR-2003  
 LOCUS  
 DEFINITION  
 OSJNEf02F19.1 OSJNEf Oryza sativa (japonica cultivar-group) cDNA  
 clone OSJNEf02F19.3, mRNA sequence.  
 ACCESSION  
 CB679186  
 VERSION  
 CB679186.1 GI:29682911  
 KEYWORDS  
 EST.  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 SOURCE  
 ORGANISM  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 726)  
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 Unpublished (2003)  
 Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu  
 PCR Primers  
 FORWARD: gta aaa cga cgg cca gtc  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 02 row: F column: 19  
 Seq primer: gga aac agc tat gac cat g.  
 Location/Qualifiers  
 1. 726  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="OSJNEf02F19"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSJNEf"  
 /note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
 XhoI; Uninfected Control"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.0952 Length: 726  
 Score: 120.50 Matches: 41  
 Percent Similarity: 44.03% Conservative: 18  
 Best Local Similarity: 30.60% Mismatches: 50  
 Query Match: 17.85% Indels: 25  
 DB: 14 Gaps: 4  
 US-09-684-215B-4 (1-132) x CB679186 (1-726)  
 QY 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla----- 27  
 Db 579 ACATCTGCTGTGTTGGTTGGTTCATCCGTCATCACTGTTCTGAAATAGCTCTCAG 520  
 QY 28 -----GlyGlnIleArgSerGlyGlyGlySerProThrValHisIleGlyPro 43  
 Db 519 TTAATTCAGTTTGGAAAGTTCGCGTGTGGCTTGAATGTGGAGTTGCTCCAGATCCA 460  
 QY 44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63  
 Db 459 ATCGCATAT-----CAGCTTAATGTCGCACTGGAATCATATATTGTCAG 415  
 QY 64 ValValGlySerAlaProAlaAlaSerLeuGly----- 74  
 Db 414 GTTCCTGGGGCAGTGTCTGCAGCAAAAGCTGCTTGTCTTCTACAGTAGGGGTTTGCT 355  
 QY 75 -----IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92  
 Db 354 GGTACCATTTGTTCTGGTGATGTCATTTGTCGCGTGGACGCGTAACCTATCAAAGGCAA 295  
 QY 93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112  
 Db 294 TCTGATCTATCGAGGGTTCTGGATGCTTATGCGGTTGGAGCAAGGTGAGCTTGACAATC 235  
 QY 113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 126  
 Db 234 CAAAGA-----GGCGCTGAATCTCTCGAGGTAACCTTG 202

```

RESULT 5
CB925602
LOCUS
DEFINITION
ABAI_22_F06.b1_A012 Abscissic acid-treated seedlings Sorghum bicolor
cDNA-clone ABAI_22_F06_A012 3', mRNA sequence.
ACCESSION
CB925602
VERSION
CB925602.1 GI:30161873
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphorbia; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 590)
Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
Eastman,A. and Pratt,L.H.
An EST database from Sorghum: ABA1-treated seedlings
Unpublished (2003)
Other ESTs: ABAI_22_F06_GL_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug3 (CGACCTGCAGCTGCAGCACA)
POLYA=Yes.
FEATURES
source
Location/Qualifiers
1..590
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="IS3620C"
/db_xref="taxon:4558"
/clone="ABAI_22_F06_A012"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Abscissic acid-treated seedlings"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. After 12 days, medium was
supplemented with 1 mM abscissic acid (ABA), while leaves
were misted with a solution of 1 mM ABA. Roots and leaves
were harvested after 3, 6, 12, and 24 hr and material from
all time points was combined prior to RNA isolation.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5'-prime
DraIII site is CACTGTGTG. 3'-prime DraIII site is
CACCATGTG). XhoI excises the cDNA insert."
ORIGIN
Alignment Scores:
Pred. No.: 0.139 Length: 590
Score: 117.50 Matches: 37
Percent Similarity: 44.19% Conservative: 20
Best Local Similarity: 28.68% Mismatches: 43
Query Match: 17.41% Indels: 29
DB: 14 Gaps: 4
US-09-684-215b-4 (1-132) x CB925602 (1-590)
QY 15 GlyPheAlaTLePzoIleGlyGlnAlaMetAlaIleAla-----Gly 28
Db 6 GGCTTGTCTATCCATCACTCACTGACTTAAATCGCTCCTCAGTAAATTCAGTTCGGA 65

```

```

QY 29 GlnIleArgSerGlyGlySerProThrValHisIleGlyProThrAlaPheIeuGly 48
Db 66 AAGTTCGTCGTCGCTTGAATCGTTCGCTCCGATCCCAATTCATAT-----119
QY 49 LeuGlyValValAspAsnGlyAsnGlyAlaAsgValGlnArgValValGlySerAla 68
Db 120 -----CAGCTTAAATGTTCCGACGGAGCTCTTATACTTAAGGTACCTGGGGCAGT 170
QY 69 ProAlaAlaSerIeuGlyIleSer-----Thr 77
Db 171 GTCTGCAGCAAAAGCAGGTCTTCTCCACCGCAGGGGTTTGTGGTAATATTGTTCTG 230
QY 78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
Db 231 GTGTATCATGTTGTCAGTGCAGCGCAACCTGTTAAGGGCAATCTGACCTGCTGAGG 290
QY 98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
Db 291 GTTCTGCATCATGTCGTCGAGATCAGGTGACCTTGACA-----332
QY 118 GlyThrArgThrGlyAsnValThrIeu 126
Db 333 ---ATCCGCGAGGCTCAGAAACCCCTT 356
RESULT 6
LOCUS
DEFINITION
CA228900 807 bp mRNA linear EST 25-SEP-2003
SCQFL3058G06.g Saccharum officinarum FL3 Saccharum officinarum
cDNA clone SCQFL3058G06 5', mRNA sequence.
ACCESSION
CA228900
VERSION
CA228900.1 GI:35290878
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE
1 (bases 1 to 807)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccc.unesp.br
Plate: 058 row: G column: 06
Seq primer: T7 Promoter Primer.
FEATURES
source
Location/Qualifiers
1..807
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQFL3058G06"
/lab_host="DH10B"
/clone_lib="Saccharum officinarum FL3"
/note="Organ: Base of developing inflorescence (5cm-long);
Vector: pSPori1; Site 1: Sali; Site 2: NotI; An
unidirectional cDNA library generated from [Base of
developing inflorescence (5cm-long)]. cDNA was prepared
from polyA+ mRNA using Superscript Plasmid System kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

```





pac82-164\_3239, genomic survey sequence.  
BZ561390  
VERSION BZ561390.1 GI:27181349  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonadaceae; Pseudomonas.  
REFERENCE 1 (bases 1 to 1033)  
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.  
TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library  
JOURNAL J. Bacteriol. (2002) In press  
COMMENT Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.  
FEATURES  
source  
1..1033  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pac82-164\_3239"  
/clone\_lib="pac82-164"  
/note="clinical isolate 2-164 Whole genomic shotgun library."  
ORIGIN  
Alignment Scores:  
Pred. No.: 3 78 Length: 1033  
Score: 106.50 Matches: 41  
Percent Similarity: 44.20% Conservatve: 20  
Best Local Similarity: 29.71% Mismatches: 56  
Query Match: 15.78% Indels: 21  
DB: 28 Gaps: 4  
US-09-684-215B-4 (1-132) x BZ561390 (1-1033)  
QY 4 SerAspAenPheGlnLeuSerGlnGlyGln-----GlyPheAlaIleProIleGly 21  
DB 713 TCCAGTATTTTCCACCGTCCGCGGCTCAATGGCCCTGTCTTTCCGATTCGATCGAT 654  
QY 22 GlnAlaMetAlaIle-AlaGlyGlnIleArg-----Se 32  
DB 653 GTCCGCTGACCGTTCGCGACCATTTGAAGAAAGCCGCAAGTCAAGTCCGCGCTGGC 594  
QY 32 rGlyGlySerProThrValHisIleGlyProThrAlaPheLeu-GlyLeuGlyValV 52  
DB 593 TGGCGGTGTGATCCAGGAAGTGAACAAGGATCTCGCGAGTCTCTCGGCTC----- 541  
QY 52 AlaAspAenGlyAsnGlyAlaArgValGlnArgValGlySerAlaProAlaAlas 72  
DB 540 -----GACAAGCTTCGCGCGCTGTGGCGCACTGTGTGAAGACGCGCCGCGGCCA 486  
QY 72 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerA 92  
DB 485 AGCGTGGCTGCGAGTGGCGATGTGATCTCTCAGCTGAAGCGCCAGTCGATCAACAGT 426  
QY 92 laThAlaMetAlaAspAlaLeuAsnGlyHisProGlyAspValIleSerVal---- 110  
DB 425 CCGCGGACCTGCGGACCTGGTGGGCAACATGAAGCCGGGGCAAGATCAACCTGGACG 366  
QY 111 -----ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124  
DB 365 TGATTCCGAACGCCAGCGCAAGTCCCTGAGCATGGCGGTAGGAGCGCTT 316  
RESULT 11

AZ933900  
LOCUS 719 bp DNA linear GSS 24-APR-2001  
DEFINITION BJ\_Ba0001B03r B. japonicum BAC library Bradyrhizobium japonicum genomic, genomic survey sequence.  
ACCESSION AZ933900  
VERSION AZ933900.1 GI:13775960  
KEYWORDS GSS.  
SOURCE Bradyrhizobium japonicum  
ORGANISM Bradyrhizobium japonicum  
REFERENCE 1 (bases 1 to 719)  
AUTHORS Tomkins, J.P., Wood, T.C., Stacey, M.G., Loh, J.T., Judd, A., Goicoechea, J.L., Stacey, G., Sadowsky, M.J. and Wing, R.A.  
TITLE A marker-dense, sequence-ready map of the Bradyrhizobium japonicum genome  
JOURNAL Genome Res. 11 (9), 1434-1440 (2001)  
MEDLINE 21376150  
PUBMED 11483585  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Class: BAC ends  
High quality sequence stop: 671.  
FEATURES  
source  
1..719  
/organism="Bradyrhizobium japonicum"  
/mol\_type="genomic DNA"  
/strain="USD110"  
/db\_xref="taxon:375"  
/lab\_host="E. coli"  
/clone\_lib="B. japonicum BAC library"  
/note="Vector: pIndigo536; Site\_1: HindIII"  
ORIGIN  
Alignment Scores:  
Pred. No.: 4 42 Length: 719  
Score: 103.50 Matches: 34  
Percent Similarity: 37.86% Conservatve: 19  
Best Local Similarity: 24.29% Mismatches: 62  
Query Match: 15.33% Indels: 25  
DB: 28 Gaps: 4  
US-09-684-215B-4 (1-132) x AZ933900 (1-719)  
QY 13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32  
DB 44 GGCATCGCTTTTCGATCTCTTGCACATGGAGCGGTGCTGCTCTCCGCAAGGCG 103  
QY 33 GlyGly-----GlySerProThrValHisIleGlyProThr 44  
DB 104 GCGCGCAAGCGGTGAAGCGTCCCTGGCTCGCGGCAAGTTCAGCGGTGACGCCGAG 163  
QY 45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64  
DB 164 ATCGCCGAAGCTCTGGCTTG-----CGTCCGCGACCGCGCGTGTTCGCGACGGTG 217  
QY 65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84  
DB 218 GTCTCGAAGCGTCCCGCAGCAAGGCGCGCTCTGAAATCTCCGATCTGATCACCAGGATC 277  
QY 85 AspGlyAlaProIleAsn----- 90  
DB 278 GAGCGCAAGACCGGTGATGATCCATGCTTCGATCGCTTCGAGCGGTTCGCGCTCCACTC 337  
QY 91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisProGlyAspValIleSerVal 110  
DB 338 GCGGTTCGCGCAGATCGACGCGCAGCGCGCAAGCGCTCAAG---CTCGCGATC 394



Qy 111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 130  
 Db 395 CGGTGGAGACCGACCGCCGCGGCGCAACAGAGCTCGTCATCATCTGCGCGTTCGCCG 454

RESULT 12  
 BZ549047/c  
 LOCUS  
 DEFINITION pasc1-60.1644.s1 pasc1-60 Pseudomonas aeruginosa genomic clone  
 pasc1-60.1644, genomic survey sequence.

ACCESSION BZ549047.1 GI:27152628  
 VERSION BZ549047.1  
 KEYWORDS GSS.  
 SOURCE Pseudomonas aeruginosa  
 ORGANISM Pseudomonas aeruginosa  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 947)  
 AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
 Burns, J.L., Kaul, R. and Olsen, M.V.  
 TITLE Whole-Genome-Sequence variation among multiple isolates of  
 Pseudomonas aeruginosa library  
 JOURNAL J. Bacteriol. (2002) In press  
 COMMENT Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun

FEATURES  
 source  
 1..947  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="1-60"  
 /db\_xref="taxon:287"  
 /clone="pasc1-60.1644"  
 /clone\_lib="pasc1-60"  
 /note="clinical isolate 1-60 Whole genomic shotgun  
 library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.57 Length: 947  
 Score: 103.50 Matches: 37  
 Percent Similarity: 44.74% Conservative: 14  
 Best Local Similarity: 32.46% Mismatches: 48  
 Query Match: 15.33% Indels: 16  
 DB: 28 Gaps: 3

US-09-684-215B-4 (1-132) x BZ549047 (1-947)

Qy 4 SerAspAsnPhcGlnLeuSerGlnGly-----GlyGlnGlyPheAlaIleProIleGly 21  
 Db 374 TCGCAGATCTTCACCGGTTCCGGCGCTTCATGGGCTTCCTTCGCCATCCGATCGAT 315

Qy 22 GlnAlaMetAlaIleAlaGlyGlnIleArgSer----- 32  
 Db 314 GTCGGATCAACGTCGCCGACCGAGTGTGAAGAAGCGGCAAGTTCAGTCGGCGTGGCTG 255

Qy 33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyValVal 52  
 Db 254 GCGGT-GGTGATCCAGAGAGTGAACAACGATCTCCCGAGTCTTCGGCTC----- 205

Qy 53 AspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 72  
 Db 204 ---GACAAGCGTCCGGCGGCTGGTGGCGCACTGGTGGAGAGCGGCCCGCGGGAAG 148

Qy 73 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92  
 Db 147 GGTGGCTGCGAGTGGCGATGTGATCAACAGCTGAACGCGCAGTCGATCAACAGCTCC 88

Qy 93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 106

Db 87 GCGGACCTGCCGACCTGGTGTGGAAACATGAAGCGGGGCGAC 46

RESULT 13  
 AQ989479  
 LOCUS  
 DEFINITION Rfc00025 Photorhabdus luminescens strain W14 M13 library  
 Photorhabdus luminescens genomic clone PLG00025, genomic survey  
 sequence.

ACCESSION AQ989479  
 VERSION AQ989479.1 GI:9648073  
 KEYWORDS GSS.  
 SOURCE Photorhabdus luminescens  
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Photorhabdus.

REFERENCE 1 (bases 1 to 736)  
 AUTHORS ffrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,  
 Daborn, P.J., Bowen, D. and Biatner, F.R.  
 TITLE A genomic sample sequence of the entomopathogenic bacterium  
 Photorhabdus luminescens W14: potential implications for virulence  
 Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

JOURNAL MEDLINE  
 COMMENT 20378633  
 10919786  
 Contact: ffrench-Constant RH  
 Department of Biology and Biochemistry  
 University of Bath  
 South Building, Bath BA2 7AY, UK  
 Tel: (44) 1225 826621  
 Fax: (44) 1225 826779  
 Email: bsrfc@bath.ac.uk  
 This is one of 2,122 random reads from the M13 library. For  
 annotation of identified clones (BLASTX, BLASTN and mapping to E.  
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic  
 Acids Res.  
 Seq primer: M13 Forward  
 Class: shotgun.

FEATURES  
 source  
 1..726  
 /organism="Photorhabdus luminescens"  
 /mol\_type="genomic DNA"  
 /strain="W14"  
 /db\_xref="taxon:29488"  
 /clone="PLG00025"  
 /dev\_stage="primary phase variant"  
 /clone\_lib="Photorhabdus luminescens strain W14 M13  
 library"  
 /note="Genomic DNA from strain W14 was size selected (1-2  
 kb) and then cloned into M13 Janus."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.02 Length: 726  
 Score: 103.00 Matches: 39  
 Percent Similarity: 46.55% Conservative: 15  
 Best Local Similarity: 33.62% Mismatches: 52  
 Query Match: 15.26% Indels: 11  
 DB: 28 Gaps: 2

US-09-684-215B-4 (1-132) x AQ989479 (1-726)

Qy 6 AsnPhcGlnLeuSerGlnGlyGln-----GlyPheAlaIleProIle 20  
 Db 179 TCATTGATTAAGTGAATGTGAAGCGCCAGAGGGTGGGTTCGCTATTCGACT 238

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 Db 239 GAACCTCGCTACTAAGATTATGCAGAACTATCCGTGACGACGAGTATCCGGGTTT 298

Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyAsn----- 57  
 Db 299 ATCGGCATAACGCAAGAGCTGCCACATATTTCGCTCTTCAACGCAATATTAAACCAG 358

```
QY 58 -----GlyAlaArgValGlnArgValValGlySerAlaProAlaSerLeuGlyIle 75
Db 359 CTTCAAGGGCTACGAGTATTTCAGGTAACCTACAAACGGCGCTCGCGAAAGACCGGTATT 418
QY 76 SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet 95
Db 419 AAGTGGGTGATCATTCATTACCGCTCATAACAAACCGCTATTCTCGGGTGAAACA 478
QY 96 AlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 111
Db 479 ATGGATCAAGTGGTGA-AATACCCCGCGAGCGTGGTCTGTCTACT 525

RESULT 14
AJ558965
LOCUS DEFINITION
AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
018_1_09_a08, mRNA sequence.
ACCESSION
VERSION
AJ558965.1 GI:31661537
KEYWORDS
SOURCE
ORGANISM
Antirrhinum majus (snapdragon)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Lamiales; Antirrhinaceae; Antirrhineae;
Antirrhinum.
REFERENCE
1 (bases 1 to 765)
Zachgo S., Stueber K., Saedler H., Sommer H. and Schwarz-Sommer, Z.
Antirrhinum EST collection
Antirrhinum EST collection
JOURNAL
COMMENT
Contact: Schwarz-Sommer Z
Molekulare Pflanzen-genetik
MPI fuer Zuechtungs-forschung
Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES
Location/Qualifiers
source
1..765
/organism="Antirrhinum majus"
/mol_type="mRNA"
/db_xref="taxon:4151"
/clone="018_1_09_a08"
/tissue_type="whole plant"
/clone_lib="Antirrhinum majus whole plant"

ORIGIN
Alignment Scores:
Pred. No.: 6.79 Length: 765
Score: 102.00 Matches: 34
Percent Similarity: 40.83% Conservative: 15
Best Local Similarity: 28.33% Mismatches: 41
Query Match: 15.11% Indels: 30
DB: 9 Gaps: 4

US-09-684-215B-4 (1-132) x AJ558965 (1-765)
QY 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGln 29
Db 311 ACCTCGCGGTGAGGATTGGATTCCTTCATCGACAGTGGTAAGATAGTGCCTCAG 370
QY 30 IleArgSerGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGly 49
Db 371 TTG-----ATCAATTATGGAAGAAGTTGTTCGAGCTGGTTA 406
QY 50 GlyValValAsp-----AsnAsnGlyAsnGlyAla 59
Db 407 AAT---GTTGACATAGCTCCTGATCTTATTCGAATCACTCATGTTCGAATGGAGCT 463
QY 60 ArgValGlnArgValValGlySerAlaProAlaSerLeuGly----- 74
Db 464 CTAATCTTGCAGGTCCTCGGAATAAGTTTCAGCAGCTAAGCGTGGCTTGTACTACACA 523
QY 75 -----IleSerThrGlyAspValIleThrAlaValAspGlyAlaPro 88
Db 524 CGGGGTTTCGTGTAATAATCGTACTTGGGGATATATTTTGGCGCTAGATGACAAACCT 593

QY 89 IleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 108
Db 584 GTGAGGAGTAGGGCGGAGCTATACAAACACTGGATGATGATTCCATCGGAGACAAAGTG 643

RESULT 15
CA480614
LOCUS DEFINITION
CA480614 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
018_1_09_a08, mRNA sequence.
ACCESSION
VERSION
CA480614.1 GI:24938284
KEYWORDS
SOURCE
ORGANISM
Antirrhinum majus (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 859)
NIN-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14289 row: k column: 14
High quality sequence stop: 662.
FEATURES
Location/Qualifiers
source
1..859
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6789351"
/tissue_type="embryo"
/lab_host="DH10B (TI-resistant)"
/clone_lib="NCI CGAP ZEmb2"
/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

ORIGIN
Alignment Scores:
Pred. No.: 11.3 Length: 859
Score: 100.50 Matches: 36
Percent Similarity: 40.30% Conservative: 18
Best Local Similarity: 26.87% Mismatches: 55
Query Match: 14.89% Indels: 25
DB: 14 Gaps: 5

US-09-684-215B-4 (1-132) x CA480614 (1-859)
QY 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle----- 26
Db 854 GCCTCTGGCACCTCAACTTCTCTTGGCCCAATGGTGATGACTTCTCGTCAGGAGCTC 795
QY 27 AlaGlyGlnIleArgSerGlyGlySerProThrValHisIle----- 41
Db 794 GTAGCTCTTCTCCAGGAGGAGGAGGAGGAGCGGTGCCATCTCTCTGCTCAAGTCAAG 735
QY 42 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgVal 61
Db 734 GGCCACATAGCAGAGCTTCTCTTGTATGTCACGGCAAT-----TTC 693
QY 62 GlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIle 81
Db -----
```

692 CCTCTCAGCTGTGGTGAAGCTGTAGCTCTCTCGGTGAGGATCTTCATGAGTAGTC 633  
 82 ThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla----- 98  
 632 AGTCAGGTCAGCGCCAGCAATGTCAGACGAGAGATGCGATGGCGAGCGGCTAACCTC 573  
 99 LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGly 118  
 572 GTAGATGGGCACAGTGTGGTGCACACCATCACCAGAGTCCATCAGTACGAGTGGTACG 513  
 119 ThrArg-----ThrGlyAsnVal 124  
 512 ACCAGAGCATACAGGGACAGACAGCTCGGTGAGCAACGTA 471

RESULT 16  
 AI597611  
 LOCUS  
 DEFINITION  
 similar to TR:008664 008664 BCL7C mRNA. ;, mRNA sequence.  
 ACCESSION  
 AI597611  
 VERSION  
 AI597611.1 GI:4606659  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 423)  
 REFERENCE  
 NCBI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute / National Institute of Neurological  
 Disorders and Stroke, Brain Tumor Genome Anatomy Project  
 (CGAP/BRGAP), Tumor Gene Index  
 UNPUBLISHED (1998)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
 Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: -40UP from Gibco.

FEATURES  
 source  
 1. 423  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2167707"  
 /tissue\_type="anaplastic oligodendroglioma"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Brn25"  
 /note="Organ: brain; Vector: pT73b-Pac (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCGAAGTGGAGCGCCGATAGGTTTTTTTTTTTTTTTTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 Library is normalized, and was constructed by Bento  
 Soares and M.Fatima Bonaldo."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.7 Length: 423  
 Score: 99.00 Matches: 35  
 Percent Similarity: 44.35% Conservative: 20  
 Best Local Similarity: 28.23% Mismatches: 43  
 Query Match: 14.67% Indels: 26  
 DB: 9 Gaps: 7

US-09-684-215B-4 (1-132) x AI597611 (1-423)

7 PheGlnLeuSerGln-----GlyGlyGlnGlyPheAlaIleProIleGlyGlnAla 23  
 39 TTTGAAAAAGCCAAAGGGCCCTGGGGCAACAGCAGCAGCGCGTTTCTCAGGGG 98  
 24 MetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThr---ValHisIleGly 42  
 99 TCAGGGGCATTTGGGCAGATGCGCTTGAGTGGGGGGCACCCTCCAAAGTCCTTTGTGCA 158  
 43 ProThrAlaPheLeuGlyLeuGlyValValaAspAsnAsnGlyAsnGlyAlaArgValGln 62  
 159 CCTGGGCTGCTCAGGACAGGT-----GGCAGTGGCTCAAAAACAGGG 203  
 63 ArgValValGlySerAlaProAlaIleSerLeuGlyIleSerThrGlyAspValIleThr 82  
 204 TAAGCTTCGGGGCCCTCACCTTCACGAGTCTGGA-----ACAGGCTCCTCCTTGTC 257  
 83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102  
 258 AGCATTTGGGGGTTCGCTGCTGTCGACAGTATG----- 293  
 103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGly-----Gly 118  
 294 CCCCCGGGA-----TCTCTTTTGGCCCCACCCGCTGGGGGCTGAGCCTTCTCAGGG 344  
 119 ThrArgThrGly 122  
 345 ACCCTTCCTGGG 356

BI350520 499 bp mRNA linear EST 26-JUL-2002  
 fr32b05.v1 zebrafish adult brain Danio rerio cDNA clone  
 IMAGE:4955073 5' similar to SW:ACT2\_FUGRU P53485 ACTIN, CYTOPLASMIC  
 2 ;, mRNA sequence.  
 BI350520  
 VERSION  
 BI350520.1 GI:15044966  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 Danio rerio (zebrafish)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 499)  
 REFERENCE  
 AUTHORS  
 Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,  
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,  
 Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,  
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.  
 WashU Zebrafish EST Project 1998  
 UNPUBLISHED (1998)  
 CONTACT: Stephen L. Johnson  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [zbratfish@watson.wustl.edu](mailto:zbratfish@watson.wustl.edu)  
 cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:  
 Matthew Clark. DNA Sequencing by: Washington University Genome  
 Sequencing Center Clone distribution: Genome Systems, St. Louis,  
 Missouri (web address: [www.genomesystems.com](http://www.genomesystems.com)) (email contact:  
[info@genomesystems.com](mailto:info@genomesystems.com)) and Research Genetics, Huntsville, Alabama  
 (web address: [www.resgen.com](http://www.resgen.com)) (email contact: [info@resgen.com](mailto:info@resgen.com)) and  
 Ressourcenzentrum Primatendatenbank, Berlin, Germany (web address:  
[www.rzpd.de](http://www.rzpd.de))  
 Seq primer: T7  
 High quality sequence stop: 301.  
 Location/Qualifiers  
 1. 499  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"

FEATURES  
 source  
 1. 499  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"

```

/lab_host="E. coli DH10B"
/clone_lib="zebrafish adult brain"
/name="Vector: pZiPlox; Site_1: Not
Original library was constructed in
/excise="The excision of the cDNA library was p
pZiPlox plasmids. Insert check was
library."

```

US-09-684-215B-4 (1-132) x BI350520 (1-499)

US-09-684-215B-4	(1-132)	X	CA473840	(1-870)
Programme Scores:				
Program No.:	22, 6			870
Score:	97, 50			35
Percent Similarity:	40, 62‡			17
Best Local Similarity:	27, 34‡			51
Query Match:	14, 44‡			25
DB:	14			5
Length:				
Matches:				
Conservative:				
Mismatches:				
Indels:				
Gaps:				

Alignment Scores:

US-09-684-215B-4	(1-132)	X	CA473840	(1-870)
Programme Scores:				
Program No.:	22, 6			870
Score:	97, 50			35
Percent Similarity:	40, 62‡			17
Best Local Similarity:	27, 34‡			51
Query Match:	14, 44‡			25
DB:	14			5
Length:				
Matches:				
Conservative:				
Mismatches:				
Indels:				
Gaps:				

US-09-684-215B-4 (1-132) x CA473840 (1-870)

16	Qy	PheAlaIleProlIeGlyGlnAlaMetAlaIle-----AlaGlyGlnIleArgSer	32
811	Db	TTTTTCTGGCCAAATGGTGTATGACCTTGTCCGTACAGGCAGCTCGTAGCTTTCTCTCAGGGA	752
33	Qy	GlyGlyGlySerProThrValHisIle-----GlyProThrAlaPheLeu	47
751	Db	GGAGGAGGAGCAGCCGGTGCCTATCTCCGTCTCAAGTCAAGGGCCACATACGAGATT	692
48	Qy	GlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySer	67
691	Db	CTCTCTGTATGTCAGGACAAT-----TTCCCTCTCAGCTGTGTGGT	650
68	Qy	AlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAla	87
649	Db	GAAGCTGTAGCGCTCTCTCGGTCCAGGATCTTCATGAGGTAGTCAGTCAGTCACGCCAC	590
88	Qy	ProfileAsnSerAlaThrAlaMetAlaAspAla-----LeuAsnGlyHisHispro	104
589	Db	CNAGTCCACGACGGAGGATGGCATGGGCGCAGGGCGTTAACCTCGTAGATGGGCACAGTGTG	530
105	Qy	GlyAspValIleIleSerValThrTrpGlnThrLysSerGlyGlyThrArg-----	120
529	Db	GGTGACACCATCACACAGATCCATCCAGTACCAGTGGTACGACCAGAGGCATACAGGA	470
121	Qy	-----ThrGlyAsnVal	124
469	Db	CAGCACAGCCCTGGATGGCAACGTA	446

CA148171

LOCUS

### DEFINITION

ACCESSION

**VERSION**

## KEYWORDS

**SOURCE**

## ORGANISMS

REFERENCE 1 (bases 1 to 444)  
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 CONTACT: Arruda P  
 Centro de Biologia Molecular e Engenhariaia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parruda@unicamp.br  
 Clone distribution: clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bccccenter.fcav.unesp.br  
 Plate: 017 row: F column: 10  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1..444  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCBRZ1017F10"  
 /lab\_host="DH10B"  
 /clone\_lib="R21"  
 /note="Organ: Shoot-root transition zone from young plants  
 (large insert library); Vector: pSport1; Site: SalI;  
 Site 2: NotI; An unidirectional cDNA library generated  
 from [Shoot-root transition zone from young plants (large  
 insert library)]. cDNA was prepared from polyA+ mRNA  
 using SuperScript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. Details  
 of each source of RNA and library construction can be  
 obtained at http://sucest.lad.ic.unicamp.br/public"

FEATURES  
 source  
 1..444  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCBRZ1017F10"  
 /lab\_host="DH10B"  
 /clone\_lib="R21"  
 /note="Organ: Shoot-root transition zone from young plants  
 (large insert library); Vector: pSport1; Site: SalI;  
 Site 2: NotI; An unidirectional cDNA library generated  
 from [Shoot-root transition zone from young plants (large  
 insert library)]. cDNA was prepared from polyA+ mRNA  
 using SuperScript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. Details  
 of each source of RNA and library construction can be  
 obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.61 Length: 444  
 Score: 97.00 Matches: 26  
 Percent Similarity: 47.63% Conservative: 14  
 Best Local Similarity: 30.95% Mismatches: 26  
 Query Match: 14.37% Indels: 18  
 DB: 13 Gaps: 2

US-09-684-215B-4 (1-132) x CAL48171 (1-444)  
 QY 54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaSerLeu 73  
 Db 32 AATGTTGGACCGAGCTTTATCTACTTAAGTACCTGGGGCGAGTGTGACCAAGCA 91  
 QY 74 GlyIleSer-----ThrGlyAspValIleThr 82  
 Db 92 GGCCTTGTCTCCGACCGGAGGGTTCGCTGTGTACATGTTCTTGGTGTGTATCGTT 151  
 QY 83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102  
 Db 152 CGAGTGGATGGCAACCTGTTAAGGCAATCAGACCTCTCAGGGTTCGTGATGACTAT 211  
 QY 103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122  
 Db 212 GCGCTCGGGGATCAGTACCTTGACA-----ATCGGGCGGAGC 250  
 QY 123 AsnValThrLeu 126  
 Db 251 TCTGAACCCCTC 262

RESULT 20  
 BI721127  
 LOCUS  
 DEFINITION  
 1031054B10.y1 C. reinhardtii CC-1690, Stress II (normalized),  
 Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.  
 BI721127  
 ACCESSION

VERSION BI721127.1 GI:15696822  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii  
 ORGANISM Chlamydomonas reinhardtii  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 543)  
 Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,  
 Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.  
 Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
 Unicellular System for Analyzing Gene Function and Regulation in  
 Vascular Plants. Project: 1031  
 Unpublished (2001)  
 JOURNAL Contact: Charles Hauser  
 COMMENT DCMB Box 91000  
 Duke University  
 Durham, NC 27708-1000  
 Tel: 919 613 8159  
 Fax: 919 613 8177  
 Email: chauser@duke.edu.  
 Location/Qualifiers  
 1..543  
 /organism="Chlamydomonas reinhardtii"  
 /mol\_type="mRNA"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_lib="C. reinhardtii CC-1690, Stress II  
 (normalized), Lambda Zap II"  
 /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:  
 XhoI; Stress condition II library, constructed by John  
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
 cells grown to mid-log phase in TAP (NH4+ - containing)  
 and shifted to TAP - NO3- (24hrs); H2 production  
 conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant  
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
 PolyA mRNA was purified from each sample, pooled and cDNA  
 synthesized. The cDNA was directionally cloned into lambda  
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')  
 sites. pBluescript II SK- plasmids were excised from the  
 lambda Zap clones by superinfection with ExAssist  
 (Stratagene) phage. The library was normalized using  
 method 4 described in Bonaldo et al., (1996) Genome  
 Research 6: 791-806."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 12.8 Length: 543  
 Score: 97.00 Matches: 31  
 Percent Similarity: 39.83% Conservative: 16  
 Best Local Similarity: 26.27% Mismatches: 49  
 Query Match: 14.37% Indels: 22  
 DB: 12 Gaps: 4

US-09-684-215B-4 (1-132) x BI721127 (1-543)  
 QY 10 SerGlnGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29  
 Db 153 GCATCTCTCCGCGTGGGCTTCGGCTCCCATCGACAGTGTGCGCGCGCTGTAGACCA 212  
 QY 30 IleArgSerGlyGlyGly-----SerProThrValHisIleGly-----ProThr 44  
 Db 213 ATCTGTACCTACCGCCCGCTGTGCGCGCGCTGTGAGTGTAGCTGCGCGCGCGAG 272  
 QY 45 AlaPheLeuGlyLeuGlyValValAlaAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64  
 Db 273 GTGCTCAAGCAGCTGGGC-----CAGCGCGCGCTGTGCTGTAGAGGTG 317  
 QY 65 ValGlySerAlaProAlaAlaSerLeuGlyIleSer-----76  
 Db 318 CCCAAGGCGTCCCCCGGAGAGAGCGCGGATCAGCCACCATGGCGACCGCTTCAGC 377  
 QY 77 -----ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92

Db 378 GGCTCCCTGGTGTGGCGGACATCATCACAGGCATTGACGGCAAGCGCGTAAGAACTAT 437  
Qy 93 ThrAlaMetAlaAspAlaLeuAenGlyHisHisProGlyAspValIleSerVal 110  
Db 438 TCGACCTGGTGGAGCGCTGGATGAGAACGGCGTGGCGACACAGTCAAGGTG 491

RESULT 21  
AW285510 546 bp mRNA linear EST 19-JUL-2000  
LOCUS LG1\_241\_E05\_g1\_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA  
DEFINITION sequence.  
ACCESSION AW285510  
VERSION AW285510.1 GI:6675354  
KEYWORDS EST.  
SOURCE Sorghum bicolor (sorghum)  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 546)  
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.  
TITLE An EST database from Sorghum: light-grown seedlings  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: T7  
High quality sequence start: 126  
High quality sequence stop: 546  
POLYA=Yes.

FEATURES  
source  
1..546  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/db\_xref="taxon:4558"  
/clone\_lib="Light Grown 1 (LG1)"  
/notes="Organ: 10- to 14-day-old light-grown (greenhouse)  
seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:  
EcoRI; The library was made from poly-A RNA in the cloning  
vector lambda ZAP II. Clones to be sequenced were  
prepared by mass excision."

ORIGIN  
Alignment Scores:  
Pred. No.: 16.2 Length: 546  
Score: 96.00 Matches: 25  
Percent Similarity: 47.82% Conservative: 15  
Best Local Similarity: 29.76% Mismatches: 26  
Query Match: 14.22% Indels: 18  
DB: 10 Gaps: 2  
US-09-684-215B-4 (1-132) x AW285510 (1-546)

Qy 54 AsnAsnGlyAsnGlyAlaArgValGlnArgValGlySerAlaProAlaAlaSerLeu 73  
Db 56 AATGTTCCCAACGGAGCTCTTATCTTAAGGTACTCTGGGGCAGTGTGCAGCCAAAGCA 115  
Qy 74 GlyIleSer-----ThrGlyAspValIleThr 82  
Db 116 GGTCTTGCTCCAACCGGCGGGGTTTCTGCTGTAATATTGTTCTGGGTGATCATCGTT 175  
Qy 83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102  
Db 176 GCAGTGGACGGCAACCTGTTAAAGGGCAAATCTGACCTGCTGAGGGTCTTGATGACTAT 235

Qy 103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122  
Db 236 GCGGTGGAGATCAGTGTGACCTTGACA-----ATCCGCGGAGGC 274  
Qy 123 AsnValThrIleu 126  
Db 275 TCAGAAACCCIT 286

RESULT 22  
AW285527 551 bp mRNA linear EST 19-JUL-2000  
LOCUS LG1\_241\_G05\_g1\_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA  
DEFINITION sequence.  
ACCESSION AW285527  
VERSION AW285527.1 GI:6675371  
KEYWORDS EST.  
SOURCE Sorghum bicolor (sorghum)  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 551)  
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.  
TITLE An EST database from Sorghum: light-grown seedlings  
JOURNAL Unpublished (2000)  
COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Sequences have been trimmed to exclude PolyA, vector and regions  
below Phred quality 16. The threshold for highest quality sequence  
is 20.  
Seq primer: T7  
High quality sequence start: 16  
High quality sequence stop: 552  
POLYA=Yes.

FEATURES  
source  
1..551  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/db\_xref="taxon:4558"  
/clone\_lib="Light Grown 1 (LG1)"  
/notes="Organ: 10- to 14-day-old light-grown (greenhouse)  
seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:  
EcoRI; The library was made from poly-A RNA in the cloning  
vector lambda ZAP II. Clones to be sequenced were  
prepared by mass excision."

ORIGIN  
Alignment Scores:  
Pred. No.: 16.5 Length: 551  
Score: 96.00 Matches: 25  
Percent Similarity: 47.62% Conservative: 15  
Best Local Similarity: 29.78% Mismatches: 26  
Query Match: 14.22% Indels: 18  
DB: 10 Gaps: 2  
US-09-684-215B-4 (1-132) x AW285527 (1-551)

Qy 54 AsnAsnGlyAsnGlyAlaArgValGlnArgValGlySerAlaProAlaAlaSerLeu 73  
Db 61 AATGTTCCCAACGGAGCTCTTATCTTAAGGTACTCTGGGGCAGTGTGCAGCCAAAGCA 120  
Qy 74 GlyIleSer-----ThrGlyAspValIleThr 82  
Db 121 GGTCTTGCTCCAACCGGCGGGGTTTCTGCTGTAATATTGTTCTGGGTGATCATCGTT 180  
Qy 83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102  
Db 181 GCAGTGGACGGCAACCTGTTAAAGGGCAAATCTGACCTGCTGAGGGTCTTGATGACTAT 240

```

QY 103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
Db 241 GCGTCGCGAGATCAGGTACCTTGACA-----ATCCGGCGGAGGC 279
QY 123 AsnValThrLeu 126
Db 280 TCAGAAACCCCT 291

RESULT 23
BU557763/c 1093 bp mRNA linear EST 16-SEP-2002
LOCUS AGNCOURT_10214312 NIH_MGC_107 Homo sapiens cDNA clone
DEFINITION IMAGE:6585799 5', mRNA sequence.
ACCESSION BU557763
VERSION BU557763.1 GI:22908059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1093)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2798 row: f column: 07
High quality sequence stop: 299.
Location/Qualifiers
1. .1093
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6585799"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/notes="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES
source
Alignment Scores:
Pred. No.: 44.2 Length: 1093
Score: 96.00 Matches: 48
Percent Similarity: 35.43% Conservatives: 14
Best Local Similarity: 27.43% Mismatches: 46
Query Match: 14.22% Indels: 68
DB: 13 Gaps: 9

US-09-684-215B-4 (1-132) x BU557763 (1-1093)

QY 23 AlaMetAlaIleAlaGlyGlnIleArgSerGlyGly-----GlySer 36
Db 789 GCTCGGGCGGATGGGAGTCAGCGCAGAAACGGGGGTGGATGGCGGGCGGCTAAGCTAT 730
QY 37 ProThrValHisIleGlyProThrAla----- 45
Db 729 CCACAGCGGCGCCCGCGGCCCCCAGGTGCAGAGCCCCCAGCGGGCGGCGGCGCTCGG 670

QY 46 -----PheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAla 59
Db 669 GCCAAGGGGGGCGACCTGGAGTAGGGGGTACGGGATTGC---GGCAATGGCTCAGCGCGCT 613
QY 60 ArgValGln-----ArgValValGly--- 66
Db 612 CGCTACACCGGTAAACCAAGTCTCCGGTGAGGATCTCTCGGGAGGATATTGGGTGC 553
QY 67 -----SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 83
Db 552 AGGAGCGCATGCCAACCGGGCGGGCGGAGTGGCGACGGGGGACGGCTACACCTG 493
QY 84 ValAsp-----GlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla--- 98
Db 492 GTAGAGTGGCACCGGGGTGGGGACCC---CGTGCCAGACGTCCATGCCCGATGCCAGG 436
QY 99 -----LeuAsnGlyHis----- 102
Db 435 GTGGCGCCAGAGGATAGAGGACAGCAGCGCTGGATGGCCAGCTACATGACCGGGG 376
QY 103 -----HisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
Db 375 GTTGAAGGTGTCAACATATATGGGAGTCA-TCTTCTCTCTGTGGCTTGGGTTTCAGGG 317
QY 118 GlyThrArgThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 316 GGACCTCGGTGACGACGACCTGGGTGCTCTCTCCGGGGCCACGGGCA 272

RESULT 24
BU5576076 1394 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_4761.Y2 msh Pseudomonas aeruginosa genomic clone msh2_4761,
DEFINITION genomic survey sequence.
ACCESSION BU5576076
VERSION BU5576076.1 GI:27211137
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1394)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence Variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) in press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1. .1394
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_4761"
/clone_lib="msh"
/notes="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Alignment Scores:
Pred. No.: 62.8 Length: 1394
Score: 96.00 Matches: 42
Percent Similarity: 37.20% Conservatives: 19
Best Local Similarity: 25.61% Mismatches: 43
Query Match: 14.22% Indels: 60
DB: 28 Gaps: 9

```

```
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/notes="vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Alignment Scores: 20.8 Length: 600
Pred. No.: 95.50 Matches: 33
Score: 45.00% Conservative: 21
Percent Similarity: 27.50% Mismatches: 45
Best Local Similarity: 14.15% Indels: 21
Query Match: 14 Gaps: 6
DB: 14

US-09-684-215B-4 (1-132) x CF787130 (1-600)
QY 12 GlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArg 31
Db 554 GGGGCAACAGGACAGGAGGCGGCTTCTCAGGGGTCAGGGGTCATTGGGCGCATGCGC 495
QY 32 SerGlyGlyGlySerProThrVal---HisIleGlyProThrAlaPheLeuGlyLeuGly 50
Db 494 TTGAGTGGGGGGGGCCCTCGTGTCTCTATCACCCTGGGCTGCTCGGGGACAGGT 435
QY 51 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValGlnArgValGlySerAlaProAla 70
Db 434 -----GGCAGCTGGCTCAAGAGCGGGTAAGCTTCAGGGGCGCTCAGCCTCC 390
QY 71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAlaAspGlyAlaProIleAsn 90
Db 389 AGCAATTCGGA-----ACAGGCTCCTCTGTGTCAGCATCGGGGGTTCATCGGTGCGG 336
QY 91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
Db 335 CCTGCA-----GGTATGCCCCCAGGA-----TCTCTC 309
QY 111 ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAlaGlyGlyPro 130
Db 308 TCCTGGCCCGAGCTGTGGGGGCTGAGCATCTCAGGG-----ACCCCTTCGGTGTGCA 255

RESULT 26
CF787130/c 600 bp mRNA linear EST 21-OCT-2003
LOCUS 854999 MARC 4PIG Sus scrofa cDNA 5', mRNA sequence.
DEFINITION CF787130.1 GI:37791691
VERSION EST.
KEYWORDS Sus scrofa (pig)
SOURCE Sus scrofa
ORGANISM Sus scrofa
REFERENCE 1 (bases 1 to 600)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J., Wray,J.B. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith RPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: TMW8001 row: 1 column: 11
Seq primer: GTAATCAGCTCACTATAGG.
Location/Qualifiers
1. 600
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"

FEATURES
source
```



Plate: 0037, row: j column: 5.  
 FEATURES  
 Location/Qualifiers  
 1. 647  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /clone="scac00371.j.05"  
 /tissue\_type="mixed"  
 /clone\_lib="Sus Scrofa library (scac)"  
 /note="vector: p773D-pac vector; tissues: adipose tissue, brain, kidney, liver, muscle, ovary, testis, heart, hypothalamus, pancreas, skin, spleen, thymus, placenta, pituitary gland, seminal vesicle, small intestine, uterus, adrenals, bulbo urethral gland, cerebral trunk, epididymis, female gonad, gall-bladder, hippocampus, large intestine, male gonad, melanocytes, stomach, udder"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 23.2 Length: 647  
 Score: 95.50 Matches: 33  
 Percent Similarity: 45.00% Conservative: 21  
 Best Local Similarity: 27.50% Mismatches: 45  
 Query Match: 14.15% Indels: 21  
 DB: 13 Gaps: 6

US-09-684-215B-4 (1-132) x BX674871 (1-647)

Qy 12 GlyGlyGlnGlyPheAlaIleProIleGlyAlaMetAlaIleAlaGlyGlnIleArg 31  
 Db 600 GGGCAACAGACAGCAGCGCGCTTCTCAGGGGTGACGGGGCATTTGGGCGAGTCGC 541

Qy 32 SerGlyGlySerProThrVal---HisIleGlyProThrAlaPheLeuGlyLeuGly 50  
 Db 540 TTGAGTGGGGGGCGCCCTCGTCTCTGTATCACCTCGGCTGCTCGGGGACAGGT 481

Qy 51 ValValAspAsnAnglyAenglyAlaArgValGlnArgValValGlySerAlaProAla 70  
 Db 480 -----GGCACTGGCTCAAGACGGGGTAAAGCTTCAGGGGGCTCAGCGCTCC 436

Qy 71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90  
 Db 435 AGCAATTCTGA-----ACAGGCTCTCTGTCAGCATCGGGGTTCATCCGTGCG 382

Qy 91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisProGlyAspValIleSerVal 110  
 Db 381 CCTGCA-----GGTATGCCCGCCAGGA-----TCTCTC 355

Qy 111 ThrTrpGlnThrLysSerGlyThrArgThrGlyAsnValThrLeuAlaGluGlyPro 130  
 Db 354 TCTGGCCCACTGTGGGGGTGAGCATCTCTCAGG-----ACCCCTTCGGTGTCTCA 301

RESULT 27  
 BI378928  
 LOCUS  
 DEFINITION  
 BFLG1.000559 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPG498) Branchiostoma floridae cDNA clone MPMPG498P1235 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Branchiostoma floridae (Florida lancelet)  
 Branchiostoma floridae  
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.

REFERENCE  
 AUTHORS  
 Herwig, R., Vengron, M. and Lehrach, H.  
 TITLE  
 New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL  
 MEDLINE  
 Genome Res. 13 (6A), 1056-1066 (2003)  
 22683279

12799346  
 CONTACT: Panopoulou G  
 Laboratory 145, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr.63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1235  
 Fax: +49 30 8413 1238  
 Email: panopoulou@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
 Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers  
 FORWARD: 5' CCCAGGCTTTACACTTTATGTTCCGGCTCG 3' (M13RSP)  
 BACKWARD: 5' GCTATTACGACGCTGGCGAAGGGGATGTG 3' (M13FSP)  
 Insert Length: 1200 Std Error: 0.00  
 Seq primer: 5'-CCGGTCCGGAATTCCTCCGGT-3', pSport3/86  
 High quality sequence stop: 677.

FEATURES  
 Location/Qualifiers  
 1. 677  
 /organism="Branchiostoma floridae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7739"  
 /clone="MPMPG498P1235"  
 /tissue\_type="whole embryo"  
 /dev\_stage="5-6 hrs (gastrula stage)"  
 /lab\_host="E.coli, XL1 blue"  
 /clone\_lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPG498)"  
 /note="vector: pSport3; Site 1: SalI, XbaI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); Oligoat primed and directionally cloned in pSport1 vector using a NotI (5'-TGACTAGTCTTATGATCGAGCGCGCC (T)15-3' and a SalI 5'-TCGACCCAGCGGTCCG-3' adapters (Gibco BRL)."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 24.8 Length: 677  
 Score: 95.50 Matches: 21  
 Percent Similarity: 61.11% Conservative: 12  
 Best Local Similarity: 38.89% Mismatches: 18  
 Query Match: 14.15% Indels: 3  
 DB: 12 Gaps: 1

US-09-684-215B-4 (1-132) x BI378928 (1-677)

Qy 58 GlyAlaArgValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThr 77  
 Db 143 GGAGTACTGGTGCACAGGATCATTTGTGGGTTCACCTGCATACAGTCGGGTATCGTCT 202

Qy 78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97  
 Db 203 GGTGACGTATCACCAGCATCAACGCTGATAGGTGACGTGACGACGAGACATCTACGAC 262

Qy 98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 111  
 Db 263 GCGGTAAATAGC-----GGGCAGCAGCTGAACATCACC 295

RESULT 28  
 CK161112  
 LOCUS  
 DEFINITION  
 FGAS042794 Triticum aestivum FGAS: TaLT5 Triticum aestivum cDNA, mRNA sequence.

ACCESSION  
 VERSION  
 CK161112  
 CK161112.1 GI:38989886

**KEYWORDS**  
 Triticum aestivum (bread wheat)  
**SOURCE ORGANISM**  
 Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooideae; Triticeae; Triticum.  
**REFERENCE**  
 1 (bases 1 to 881)  
**AUTHORS**  
 Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,  
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,  
 Linke, M.G., McCarthy, B.L., Monroy, A., Muzak, I., Nilsson, D.,  
 Penniket, C., Roach, J.L. and Sarhan, F.  
**TITLE**  
 Functional Genomics of Abiotic Stress In Wheat and Canola Crops  
**JOURNAL**  
 Unpublished (2003)  
**COMMENT**  
 Contact: Wm L Crosby  
 Bioinformatics  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: fgas\_este@cs.usask.ca  
 This sequence is the direct result of the Base calling software  
 Pired (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [123,692].  
 Plate: TaLT547 row: K column: 14.  
**FEATURES**  
 Location/Qualifiers  
 1..881  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Wheat line PI 178383"  
 /db\_xref="taxon:4565"  
 /lab\_host="DH5 alpha"  
 /clone\_lib="Triticum aestivum FGAS: TaLT5"  
 /note="Organ: Crown; Vector: pGEM-T; SSH (suppression  
 subtractive hybridization) cDNA library from genotype  
 PI178383 cold hardened at 2 C for 21 days and 49 days  
 (equal amount of cDNA pooled together before subtraction,  
 tester) and subtracted against genotype Norstar cold  
 hardened at 2 C for 1 day (24 H) (driver). Modified Smart  
 cDNA (Clontech) priming and non-directional cloning"  
**ORIGIN**  
 Alignment Scores:  
 Pred. NO.: 40.6 Length: 881  
 Score: 95.00 Matches: 35  
 Percent Similarity: 44.88% Conservative: 22  
 Best Local Similarity: 27.56% Mismatches: 58  
 Query Match: 14.07% Indels: 12  
 DB: 14 Gaps: 6  
 US-09-684-215B-4 (1-132) x CK161112 (1-881)  
 Qy 7 PheGlnLeuSerGInGly--GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAla 25  
 Db 509 TTTGGCATCGGCTCTGGTTTGGGCTCTGGCTTCGGCATCGGCTTCGGCTTCGGCTTCGGC 569  
 Qy 26 IleAlaGlyGlnIleArgSerGlyGlyGlySerPro---ThrValHisIleGlyProThr 44  
 Db 569 ATCGGCTCAGGCTTGGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCG 625  
 Qy 45 AlaPheLeuGlyLeuGlyValValAlaSpAsnAndGlyAsnGlyAlaArgValGlnArgVal 64  
 Db 626 GGCTTCATTCGCTCGGGTTTGGGCTCTGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCG 685  
 Qy 65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84  
 Db 686 TTGGGCTCANGCTTAGGCTCAGGTTTGGGCATTGGGTAAAGC-----TTGGGCTCCGGC 739  
 Qy 85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 104  
 Db 740 TTCGGTTTCGGGTTTGGCATCGGCTCAGGCTCAGGCTTCGGGCTCTGGGTTGATC---CACTGGCGC 796



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 20:30:12 ; Search time 375.474 Seconds  
(without alignments)  
1493.479 Million cell updates/sec

Title: US-09-684-215B-4

Perfect score: 675  
Sequence: 1 TAASDNFQLSQGGQFAIPI.....QTKSGGTRGNVTLAEGPPA 132

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09684215/runat\_29042004\_061304\_13179/app\_query.fasta\_1.1180  
-DB=N Geneseq\_29Jan04 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=30  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215@cgn\_1\_1\_819 -runat\_29042004\_061304\_13179 -NCPU=6 -ICPU=3  
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_29Jan04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	675	100.0	396	4 AAL40769	Aal40769 Nucleotid
2	675	100.0	672	4 AAL40772	Aal40772 Nucleotid
3	675	100.0	702	2 AAZ20206	Aaz20206 Mycobacte
4	675	100.0	702	4 AAL40770	Aal40770 Nucleotid
5	675	100.0	702	6 ABK14140	Abk14140 DNA encod
6	675	100.0	1002	6 AAD47078	Aad47078 Mycobacte
7	675	100.0	1002	6 AAD47077	Aad47077 Mycobacte
8	675	100.0	1002	6 AAD28336	Aad28336 Mycobacte

9	675	100.0	1002	6 AAD28337	Aad28337 Mycobacte
10	675	100.0	1068	2 AAX34251	Aax34251 Mycobacte
11	675	100.0	1143	2 AAX34252	Aax34252 Mycobacte
12	675	100.0	1629	5 AAS03793	Aas03793 M. tuberc
13	675	100.0	1742	4 AAL40771	Aal40771 Nucleotid
14	675	100.0	1871	4 AAL40768	Aal40768 Nucleotid
15	675	100.0	2181	8 ADA26353	Ada26353 Mycobacte
16	675	100.0	2190	6 AAD47084	Aad47084 Mycobacte
17	675	100.0	2190	6 AAD28343	Aad28343 Mycobacte
18	675	100.0	2191	4 AAL40773	Aal40773 Nucleotid
19	675	100.0	2286	6 ABK14128	Abk14128 DNA encod
20	675	100.0	2287	2 AAZ20194	Aaz20194 Mycobacte
21	675	100.0	2287	6 AAD47083	Aad47083 Mycobacte
22	675	100.0	2287	6 AAD28342	Aad28342 Mycobacte
23	675	100.0	2451	8 ADA26360	Ada26360 Mycobacte
24	675	100.0	2487	8 ADA26359	Ada26359 Mycobacte
25	675	100.0	2637	8 ADA26358	Ada26358 Mycobacte
26	675	100.0	2808	6 AAD47110	Aad47110 Mycobacte
27	675	100.0	2808	8 ADA26357	Ada26357 Mycobacte
28	675	100.0	3030	8 ADA26355	Ada26355 Mycobacte
29	675	100.0	3060	8 ADA26363	Ada26363 M. bovis
30	675	100.0	3104	8 ADA26362	Ada26362 Mycobacte
31	675	100.0	3474	8 ADA26361	Ada26361 Mycobacte
32	675	100.0	110000	4 AAI99682_01	Continuation (2 of
33	675	100.0	110000	4 AAI99683_01	Continuation (2 of
34	670	99.3	447	2 AAT91403	Aat91403 Mycobacte
35	670	99.3	447	2 AAT91466	Aat91466 Mycobacte
36	670	99.3	447	2 AAV44342	Aav44342 Mycobacte
37	670	99.3	447	2 AAV64450	Aav64450 M. tuberc
38	670	99.3	447	2 AAZ19040	Aaz19040 M. tuberc
39	670	99.3	447	2 AAZ19252	Aaz19252 M. tuberc
40	670	99.3	447	5 AAS03780	Aas03780 M. tuberc
41	670	99.3	447	6 AAD47080	Aad47080 Mycobacte
42	670	99.3	447	6 AAD28339	Aad28339 Mycobacte
43	670	99.3	1872	2 AAT91414	Aat91414 Mycobacte
44	670	99.3	1872	2 AAT91477	Aat91477 Mycobacte
45	670	99.3	1872	2 AAV44355	Aav44355 Mycobacte

ALIGNMENTS

RESULT 1  
AAL40769  
ID AAL40769 standard; DNA; 396 BP.

XX  
AC AAL40769;  
XX  
DT 03-OCT-2002 (first entry)  
XX  
DE Nucleotide sequence encoding Ral2 protein.

XX  
KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.  
XX  
OS Unidentified.

XX  
FH Key Location/Qualifiers  
CDS 1..396  
FT /\*tag= a  
FT /product= "Ral2 protein"  
FT /note= "No start or stop codon"  
XX  
FN WO200125401-A2.

XX  
PD 12-APR-2001.

XX  
PF 06-OCT-2000; 2000WO-US027652.

XX  
PR 07-OCT-1999; 99US-0158585P.

XX  
PA (CORI-) CORIXA CORP.

PI Skeiky Y, Guderian J;  
 XX WPI; 2001-266299/27.  
 DR P-PSDB; AAO22138.  
 XX  
 PT Recombinant nucleic acid molecule for producing high yield expression of  
 PT desired fusion polypeptides, encodes fusion polypeptide comprising  
 PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.  
 XX  
 XX Claim 1; Fig 2; 39pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2 protein  
 XX  
 SQ Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7,988-55 Length: 396  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-684-215B-4 (1-132) x AAL40769 (1-396)  
 QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 Db 1 ACGGCGCGGTCCGATTAATTCAGCTGTCAGGCTGGGCGAGGATTCGCGATTCGCGATC 60  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 61 GGGCAGGCGATGGCGATTCGGGCGCAGATCCGATCGGCTGGGGGTGTCACCCACCGTTTCA 120  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArg 60  
 Db 121 ATCGGCGCTACCGCTTCTCGCTTGGGTGTGTTCGACAAACACGCGACGCGCGACGA 180  
 QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 181 GTCCACCGGTGTTCGGAGCGCTCCGGCGGCAAGTCTCGGATCTCCACCGCGGACGTG 240  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 241 ATCCCGCGGTTCACGGCGCTCCGATCAATCGGCGCACCGGATGGCGGACGCGCTTAAC 300  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyIleThrArg 120  
 Db 301 GGGCATCAPCCCGGTGACGTATCTCGGTGACCTCGGCAACCAAGTCGGGCGGACGCGT 360  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 361 ACAGGGGAACGTGACATTCGCGCGGAGGACCCCGCGCC 396  
 RESULT 2  
 AAL40772  
 ID AAL40772 standard; DNA; 672 BP.  
 XX  
 XX AAL40772;  
 AC

XX 06-AUG-2003 (revised)  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Nucleotide sequence encoding Ral2-mammaglobin fusion protein.  
 XX  
 KW Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
 KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
 KW immunogen; cytokine; gene; ds.  
 XX  
 OS Mammalia.  
 OS Unidentified.  
 OS Chimeric.  
 FH  
 FH Key Location/Qualifiers  
 FT 4..666  
 FT CDS /\*tag= a "Ral2-mammaglobin fusion protein"  
 FT /product= "Ral2-mammaglobin fusion protein"  
 XX  
 PN WO200125401-A2.  
 XX  
 XX 12-APR-2001.  
 XX  
 XX 06-OCT-2000; 2000WO-US027652.  
 XX  
 XX 07-OCT-1999; 99US-0158585P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Skeiky Y, Guderian J;  
 PI  
 XX  
 XX WPI; 2001-266299/27.  
 DR P-PSDB; AAO22141.  
 XX  
 PT Recombinant nucleic acid molecule for producing high yield expression of  
 PT desired fusion polypeptides, encodes fusion polypeptide comprising  
 PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.  
 XX  
 PS Disclosure; Fig 5; 39pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ral2-mammaglobin fusion protein. (Updated  
 CC on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,458-54 Length: 672  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-684-215B-4 (1-132) x AAL40772 (1-672)  
 QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 Db 25 ACGGCGCGGTCCGATTAATTCAGCTGTCAGGCTGGGCGAGGATTCGCGATTCGCGATC 84

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCCACCACCGTTCAT 144  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 145 ATCCGGGCTACCGCTTCTCGGTGGTGTGTCGACACACGCGACGCGACGCA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACCGCGTCCGACGGCTTCCGATCACTCCGCCACCGGATGGCGGCGGCTTAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAAAACCAAGTCGGCGGCGACGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
Db 385 ACAGGAAACGTGACATTCGCCGAGGACCCCGGCC 420  
RESULT 3  
AAZ20206  
ID AAZ20206 standard; DNA; 702 BP.  
XX  
AC AAZ20206;  
AT 17-JAN-2000 (first entry)  
DE Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.  
KW Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD; diagnosis;  
KW therapy; vaccine; immunogen; ss.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN WO9951748-A2.  
XX  
PD 14-OCT-1999.  
XX  
PF 07-APR-1999; 99WO-US007717.  
XX  
PR 07-APR-1999; 98US-00056556.  
PR 30-DEC-1998; 98US-00223040.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skelky YAW, Alderson M, Campos-Neto A;  
XX  
DR WPI; 1999-601610/51.  
DR P-PSDB; AAY32071.  
XX  
PT New fusion proteins useful for diagnosis, prevention and treatment of  
PT tuberculosis.  
XX  
PS Example; Fig 13A-B; 83pp; English.  
XX  
CC This DNA sequence includes a coding region for a recombinant  
CC Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),  
CC termed Mtb24, composed of the antigens Ra12 and DPPD. The DNA is useful  
CC for the recombinant production of the fusion protein. Coding sequences  
CC for the antigens were modified by PCR in order to facilitate their fusion  
CC and subsequent expression of the fusion protein, and then ligated. The  
CC invention provides fusion proteins (see AAY32059-71) containing at least  
CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides  
CC encoding them are useful as vaccines for preventing tuberculosis  
CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests  
CC for detection of anti-M. tuberculosis antibodies), monitoring of disease  
CC progression, and treatment of tuberculosis. They are more effective

CC immunogens than mixtures of the individual protein components  
XX  
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.53e-54 Length: 702  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
US-09-684-215B-4 (1-132) x AAZ20206 (1-702)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 25 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGCGAGGATTCCGCATTCCGATC 84  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCCACCACCGTTCAT 144  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 145 ATCCGGGCTACCGCTTCTCGGTGGTGTGTCGACACACGCGACGCGACGCA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 264  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACCGCGTCCGACGGCTTCCGATCACTCCGCCACCGGATGGCGGCGGCTTAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAAAACCAAGTCGGCGGCGACGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
Db 385 ACAGGAAACGTGACATTCGCCGAGGACCCCGGCC 420  
RESULT 4  
AAZ40770  
ID AAZ40770 standard; DNA; 702 BP.  
XX  
AC AAZ40770;  
XX  
DT 03-OCT-2002 (first entry)  
XX  
DE Nucleotide sequence encoding Ra12-DPPD fusion protein.  
XX  
KW Ra12; serine protease antigen; Mtb32A; Mycobacterium tuberculosis;  
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
KW immunogen; cytokine; gene; ds.  
XX  
OS Unidentified.  
OS Chimeric.  
XX  
PH Key Location/Qualifiers  
FT CDS 4..696  
FT /\*tag= a  
FT /product= "Ra12-DPPD fusion protein"  
XX  
PN WO200125401-A2.  
XX  
PD 12-APR-2001.  
XX  
XX 06-OCT-2000; 2000WO-US027652.  
PF  
PR 07-OCT-1999; 99US-0158585P.  
XX  
XX (CORI-) CORIXA CORP.  
PA



CC response of Mycobacterium tuberculosis. Polypeptides of the invention are  
CC useful for diagnosing, treating or preventing M. tuberculosis infection,  
CC particularly tuberculosis infection. In particular, the polypeptides are  
CC useful as a vaccine formulation with an adjuvant to afford long-term  
CC protection in animals against the development of tuberculosis. The  
CC protein coding sequence may be used to encode a protein product for use  
CC as an immunogen to induce and/or enhance an immune response to M.  
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis  
CC fusion protein of the invention. This polynucleotide encodes 3 different  
CC proteins, each in a different reading frame. (Updated on 29-AUG-2003 to  
CC standardise OS field)

XX SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1,53e-54 Length: 702  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-4 (1-132) x ABK14140 (1-702)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
DB 25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCGGATC 84  
QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 85 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTTCA 144  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 145 ATCGGGCCCTACCGCTTCTCGGCTTGGTGTGTCGACACACGCGCAACGGCGCACGA 204  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 205 GTCCAAACGGCTGTGTCGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGCACGTG 264  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 265 ATCACCGCGTGCAGCGCTCCGATCACTCGGCCACCGCGATGCGGACGCGTTTAA 324  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 325 GGGCATCATCCGCGTACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGT 384  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
DB 385 ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC 420

## RESULT 6

ID AAD47078 standard; DNA; 1002 BP.

XX AC AAD47078;

XX DT 27-JAN-2003 (first entry)

XX MYcobacterium tuberculosis Ra35FLMUTSA mutant antigenic protein DNA.

XX vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMUTSA; mutant;  
XX gene; antigen; ds.

XX OS MYcobacterium tuberculosis.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT 4,996

XX FT /\*tag= a

XX FT /product= "Ra35FLMUTSA mutant antigenic protein"

PN WC200272792-A2.

XX 19-SEP-2002.

XX 13-MAR-2002; 2002WO-US008223.

XX 13-MAR-2001; 2001US-0275837P.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Brannon M, Guderian J;

XX WPI; 2002-759844/82.

XX P-PSDB; AAE29703.

XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,  
XX M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity  
XX against pathogenic microorganisms e.g. Leishmania and Mycobacterium  
XX tuberculosis.

XX Disclosure; Page 80-81; 155pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
XX fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
XX polynucleotide sequence encoding an antigen or an antigenic fragment from  
XX Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
XX polypeptide or its fragment. The Leishmania polynucleotide is selected  
XX from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
XX are used in methods for eliciting immune response in mammals. They are  
XX useful as vaccines to elicit protective immunity against pathogenic  
XX microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
XX polypeptides are used for enhancing the expression of polynucleotides, as  
XX in vivo diagnostic agents and for raising antibodies in a non-human  
XX animal. The invention is used in gene therapy. The present sequence is M.  
XX tuberculosis Ra35FLMUTSA mutant antigenic protein encoding DNA

SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,29e-54 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-4 (1-132) x AAD47078 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
DB 598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGATTCGGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 658 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTTCA 717  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
DB 718 ATCGGGCCCTACCGCTTCTCGGCTTGGTGTGTCGACACACGCGCAACGGCGCACGA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 778 GTCCAAACGGTGTGTCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 837  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 838 ATCACCGCGTGCAGCGCTCCGATCACTCGGCCACCGCGATGCGGACGCGTTTAA 897  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
DB 898 GGGCATCATCCGCGTACGCTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGCACGCGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132



Db 958 ACAGGGAACGTGACATGGCCGAGGACCCCGGCC 993  
RESULT 7  
AAD47077  
XX AAD47077 standard; DNA; 1002 BP.  
XX AC  
XX AAD47077;  
XX 27-JAN-2003 (first entry)  
XX Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.  
XX DE  
XX Mycobacterium tuberculosis mature Ra35 antigen; gene therapy; Ra35 antigen; gene;  
XX KW ds.  
XX KW vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen; gene;  
XX OS Mycobacterium tuberculosis.  
XX XX  
XX FH Key Location/Qualifiers  
XX CDS 4..996  
XX FT /\*tag= a  
XX FT /product= "Ra35 mature antigenic protein"  
XX FT /transl\_except= (pos:547..549, aa:Asp)  
XX FT /transl\_except= (pos:550..552, aa:Ser)  
XX XX  
XX WO200272792-A2.  
XX EN  
XX PD 19-SEP-2002.  
XX PF 13-MAR-2002; 2002WO-US008223.  
XX PR 13-MAR-2001; 2001US-0275837P.  
XX XX (CORI-) CORIXA CORP.  
XX XX Skeiky Y, Brannon M, Guderian J;  
XX FI WPI; 2002-759844/82.  
XX DR P-PSDB; AAE29702.  
XX XX  
XX FT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,  
XX FT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity  
XX FT against pathogenic microorganisms e.g. Leishmania and Mycobacterium  
XX FT tuberculosis.  
XX XX  
XX PS Disclosure; Page 79; 155pp; English.  
XX XX  
XX CC The invention relates to a recombinant nucleic acid molecule encoding a  
XX CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
XX CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
XX CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
XX CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
XX CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
XX CC are used in methods for eliciting immune response in mammals. They are  
XX CC useful as vaccines to elicit protective immunity against pathogenic  
XX CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
XX CC polypeptides are used for enhancing the expression of polynucleotides, as  
XX CC in vivo diagnostic agents and for raising antibodies in a non-human  
XX CC animal. The invention is used in gene therapy. The present sequence is M.  
XX CC tuberculosis mature Ra35 (N-terminus of MTB32A; Ra35FL) antigen encoding  
XX CC DNA  
XX XX  
XX SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2.29e-54 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215B-4 (1-132) x AAD47077 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
598 ACAGCGCGGTCCGATAAATCCAGCTGTCCAGGGTGGCAGGATTCCGCATTCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
658 GGGCAGCGGATGGCGATCGCGGCCAGATCGGATCGGGTGGGGGTCCACCACGTTTCA 717  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArg 60  
718 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTGTCGACAAACACGCAACGGCGGCACGA 777  
DB 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
778 GTCCAAACGCGTGGTGGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGGACGTG 837  
QY 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
838 ATCACCAGCGGTGACGCGGCTCCGATCACTCGCCACCGCGATGGCGGCGCTTAAC 897  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
898 GGGCATCATCCCGGTGACGTCTCTCGTGACCTCGGCAACCAAGTCGGGCGGACGCGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
958 ACAGGGAACGTGACATTCGCCGAGGACCCCGGCC 993  
RESULT 8  
AAD28336  
ID AAD28336 standard; cDNA; 1002 BP.  
XX AC  
XX AAD28336;  
XX DT 22-APR-2002 (first entry)  
XX DE Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.  
XX KW Fusion protein; antigen; serological sensitivity; immune response;  
XX KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.  
XX OS Mycobacterium sp.  
XX XX  
XX FH Key Location/Qualifiers  
XX CDS 4..996  
XX FT /\*tag= a  
XX FT /product= "Ra35 mature protein"  
XX FT /transl\_except= (pos:547..549, aa:Asp)  
XX FT /transl\_except= (pos:550..552, aa:Ser)  
XX XX  
XX PN WO200198460-A2.  
XX PD 27-DEC-2001.  
XX XX 20-JUN-2001; 2001WO-US019959.  
XX XX 20-JUN-2000; 2000US-00597796.  
XX PR 01-FEB-2001; 2001US-0265737P.  
XX XX (CORI-) CORIXA CORP.  
XX PA  
XX PI Skeiky Y, Reed S, Alderson M;  
XX XX WPI; 2002-147798/19.  
XX DR P-PSDB; AAE17566.  
XX XX  
XX PT Composition comprising MTB39 antigen and MTB32A antigen from  
XX PT Mycobacterium species, useful for eliciting immune response in a subject.  
XX XX Disclosure; Page 95; 136pp; English.  
XX XX  
XX CC The present invention relates to fusion proteins containing at least two

CC Mycobacterium species antigens, nucleotides encoding them and  
CC compositions comprising such fusion proteins. The present invention  
CC particularly relates to nucleic acids encoding fusion proteins that  
CC include two or more individual M. tuberculosis antigens which increase  
CC the serological sensitivity of sera from individuals infected with  
CC tuberculosis and methods for their use in diagnosis, prevention and  
CC treatment of tuberculosis infection. Sequences of the invention are  
CC useful for eliciting an immune response in a mammal, e.g., human,  
CC immunised with BCG. They are useful in the diagnosis, treatment and  
CC prevention of Mycobacterium infection. The fusion proteins and the  
CC polynucleotides are useful as diagnostic tools in patients infected with  
CC Mycobacterium, in vitro and in vivo assays for detecting humoral  
CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
CC diagnosis of an infection or monitoring of disease progression, as  
CC immunogens to generate or elicit a protective immune response in a  
CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
CC animal. Sequences of the invention are also used as vaccines. MTB32A  
CC fusion proteins of the invention are useful as in vivo diagnostic agents  
CC for intradermal skin test. The present sequence is a cDNA encoding  
CC Mycobacterium species MTB32A (Ra32FL) mature protein  
XX  
SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,29e-54 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215B-4 (1-132) x AAD28336 (1-1002)  
QY 1 ThrAlaAlaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 598 ACGGCGCGTCCGATACCTCCAGCTGCTCCAGGGTGGGAGGATTCGCGATTCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 658 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCACCGTTCAT 717  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 718 ATCGGGCTACCGCTCTCCGGTGGGTGGTGTGACACACGCGACGCGACGCA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 778 GTCCACGCGGTGTCGGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGACGTG 837  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 838 ATCACCAGGTGCGAGCGGCTCCGATCACTCGGCCACCGGATGGCGGACGGCTTAC 897  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLySerGlyGlyThrArg 120  
Db 898 GGGCATCATCCGCTGACGTCATCTCGTGACCTGGCAACCAAGTCCGGCGCGACGCGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
Db 958 ACAGGGAACGTGACATTGGCGGAGGAGCCCCCGGCC 993

RESULT 9  
AAD28337  
ID AAD28337 standard; cDNA; 1002 BP.  
XX  
AC AAD28337;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Mycobacterium species Ra32FLMutSA mutant cDNA.  
XX  
KW Fusion protein; antigen; serological sensitivity; immune response;  
tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.

XX OS Mycobacterium sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 4..996  
FT /\*tag= a  
FT /product= "Ra32FLMutSA protein"  
XX  
XX WO200198460-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 20-JUN-2001; 2001WO-US019959.  
XX  
XX 20-JUN-2000; 2000US-00597796.  
XX 01-FEB-2001; 2001US-0265737P.  
XX (CORI-) CORIAX CORP.  
XX  
XX Skeiky Y, Reed S, Alderson M;  
XX  
XX WPI: 2002-147798/19.  
XX P-PSDB; AAE17567.  
XX  
XX Composition comprising MTB39 antigen and MTB32A antigen from  
XX Mycobacterium species, useful for eliciting immune response in a subject.  
XX  
XX Claim 67; Page 96-97; 136pp; English.  
XX  
XX The present invention relates to fusion proteins containing at least two  
XX Mycobacterium species antigens, nucleotides encoding them and  
XX compositions comprising such fusion proteins. The present invention  
XX particularly relates to nucleic acids encoding fusion proteins that  
XX include two or more individual M. tuberculosis antigens which increase  
XX the serological sensitivity of sera from individuals infected with  
XX tuberculosis and methods for their use in diagnosis, prevention and  
XX treatment of tuberculosis infection. Sequences of the invention are  
XX useful for eliciting an immune response in a mammal, e.g., human,  
XX immunised with BCG. They are useful in the diagnosis, treatment and  
XX prevention of Mycobacterium infection. The fusion proteins and the  
XX polynucleotides are useful as diagnostic tools in patients infected with  
XX Mycobacterium, in vitro and in vivo assays for detecting humoral  
XX antibodies or cell-mediated immunity against M. tuberculosis, for the  
XX diagnosis of an infection or monitoring of disease progression, as  
XX immunogens to generate or elicit a protective immune response in a  
XX patient and for raising anti-M. tuberculosis antibodies in a non-human  
XX animal. Sequences of the invention are also used as vaccines. MTB32A  
XX fusion proteins of the invention are useful as in vivo diagnostic agents  
XX for intradermal skin test. The present sequence is a cDNA encoding  
XX Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA  
XX  
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,29e-54 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-684-215B-4 (1-132) x AAD28337 (1-1002)  
QY 1 ThrAlaAlaSerAspAsnPhleGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 598 ACGGCGCGTCCGATACCTCCAGCTGCTCCAGGGTGGGAGGATTCGCGATTCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 658 GGGCAGCGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCACCGTTCAT 717  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60

```

Db      718 ATCGGGCTACCGCTTCCTCGGTTGGTGGTTGTGCAGCAACAAACGGCAACGGCGCACGA 777
Qy      61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db      778 GTCCAAACGCGTGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 837
Qy      81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db      838 ATCACCGCGTGTGACGGCGCTCCGATCACTCGGCCACCGGATGGCGACGGCGCTTAAC 897
Qy      101 GlyHisProGlyAspValIleSerValThrTropGlnThrIlySerGlyGlyThrArg 120
Db      898 GGGCATCATCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGCGCGCACGGCGT 957
Qy      121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db      958 ACAGGGAACGTGACATTGGCCGAGGACCCCGCGCC 993

RESULT 10
AAX34251
ID AAX34251 standard; DNA; 1068 BP.
AC AAX34251;
XX
XX DT 06-JUL-1999 (first entry)
XX
XX Mycobacterium species nucleic acid sequence 50D.
XX Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
XX hybridisation; detection; vaccine; immunisation; infection; ss.
XX Mycobacterium sp.
XX WO9909186-A2.
XX
XX PD 25-FEB-1999.
XX
XX PF 14-AUG-1998; 98WO-FR001813.
XX
XX PR 14-AUG-1997; 97FR-00010404.
XX
XX PR 11-SEP-1997; 97FR-00011325.
XX
XX PA (INSP ) INST PASTEUR.
XX
XX PI Gicquel B, Portnoie D, Lim E, Pellicic V, Guigueno A;
XX PI Goguet De La Salmoniere Y;
XX
XX DR WPI; 1999-181045/15.
XX
XX DR P-PSDB; AAY05000.
XX
XX PT Mycobacterial DNA vectors containing reporter constructs - for
XX PT identifying coding or promoter sequences involved in infection-associated
XX PT protein expression.
XX
XX PS Claim 22; Fig 50D; 309pp; French.
XX
XX CC Sequences AAX34001-X34252 represent nucleic acids encoding secreted
XX CC proteins from various Mycobacterium species microorganisms. The
XX CC nucleotide sequences can be used as primers and probes for methods for
XX CC detecting and identifying mycobacteria, especially belonging to the M.
XX CC tuberculosis complex. The encoded proteins can be used in vaccines for
XX CC immunisation against a bacterial or viral infection
XX
XX SQ Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,46e-54 Length: 1068
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

```

```

US-09-684-215B-4 (1-132) x AAX34251 (1-1068)
Qy      1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db      670 ACGCGCGGTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGATTGGCCATTCGGATC 729
Qy      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db      730 GGGCAAGCGATGGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 789
Qy      41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db      790 ATCGGGCGCTACCGCTTCCTCGGCTTGGGTGTGTGTCGACAAACAAACGGCAACGGCGCACGA 849
Qy      61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db      850 GTCCAAACGCGTGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCACGTG 909
Qy      81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db      910 ATCACCGCGGTGACGGCGCTCCGATCACTCGGCCACCGGATGGCGACGGCGCTTAAC 969
Qy      101 GlyHisProGlyAspValIleSerValThrTropGlnThrIlySerGlyGlyThrArg 120
Db      970 GGGCATCATCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGCGCGCACGGCGT 1029

RESULT 11
AAX34252
ID AAX34252 standard; DNA; 1143 BP.
XX
XX AC AAX34252;
XX
XX DT 06-JUL-1999 (first entry)
XX
XX DE Mycobacterium species nucleic acid sequence 50F.
XX
XX KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
XX KW hybridisation; detection; vaccine; immunisation; infection; ss.
XX
XX OS Mycobacterium sp.
XX
XX PN WO9909186-A2.
XX
XX PD 25-FEB-1999.
XX
XX PF 14-AUG-1998; 98WO-FR001813.
XX
XX PR 14-AUG-1997; 97FR-00010404.
XX
XX PR 11-SEP-1997; 97FR-00011325.
XX
XX PA (INSP ) INST PASTEUR.
XX
XX PI Gicquel B, Portnoie D, Lim E, Pellicic V, Guigueno A;
XX PI Goguet De La Salmoniere Y;
XX
XX DR WPI; 1999-181045/15.
XX
XX DR P-PSDB; AAY04830.
XX
XX PT Mycobacterial DNA vectors containing reporter constructs - for
XX PT identifying coding or promoter sequences involved in infection-associated
XX PT protein expression.
XX
XX PS Claim 22; Fig 50F; 309pp; French.
XX
XX CC Sequences AAX34001-X34252 represent nucleic acids encoding secreted
XX CC proteins from various Mycobacterium species microorganisms. The
XX CC nucleotide sequences can be used as primers and probes for methods for
XX CC detecting and identifying mycobacteria, especially belonging to the M.
XX CC tuberculosis complex. The encoded proteins can be used in vaccines for

```

CC immunisation against a bacterial or viral infection  
 XX Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,66e-54 Length: 1143  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-684-215B-4 (1-132) x AAX34252 (1-1143)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 745 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGAGGATTCGCCATTCGGATC 804  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 DB 805 GGGCAGCGGATGGCGATCGCGGCCAGATCCGATCGGGTGGGGTCCACCACCGTTTCA 864  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
 DB 865 ATCGGGCTTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACGCGACGCGCAGCA 924  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 925 GTCCACACGCGTGTGCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGCAGTG 984  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 985 ATCACCGCGTGTGCGGAGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAA 1044  
 QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleValSerGlyGlyThrArg 120  
 DB 1045 GGGCATCATCCGCTGAGCTATCTCGGTGACCTTGGCAACCAAGTGGGCGCGCAGCGT 1104

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

DB 1105 ACAGGGAAGTGTGACATTTGGCGGAGGAGCCCGCGCC 1140

RESULT 12

AAS03793  
 ID AAS03793 standard; DNA; 1629 BP.

XX AAS03793;

AC AAS03793;

DT 29-AUG-2001 (first entry)

DE M. tuberculosis DNA encoding antigen HTCC#1 fusion protein #4.

XX TBRA12-HTCC#1; antigen; vaccine; tuberculosis; AIDS;

KW acquired immunodeficiency disease; His Tag; ds.

XX Mycobacterium tuberculosis.

OS Synthetic.

XX Key

FT Location/Qualifiers

FT 1..1629

FT /tag= a

FT /product= "TBRA12-HTCC#1"

FT /transl\_except= (pos:1621..1623,aa:Xaa)

FT /note= "Xaa= In frame STOP codon"

FT /partial

FT /note= "No start or stop codon. Although the sequence does contain an in frame stop codon, 2 further amino acids are shown in Figure 8 as being encoded by the present sequence, without a further stop codon"

FT 25..426

FT /tag= b

FT /note= "Region derived from TBRA12"

FT 427..444

FT /tag= c  
 FT /note= "Region derived from Thrombin"  
 FT 445..1629  
 FT /tag= d  
 FT /note= "Region derived from HTCC#1"

XX WO200124820-A1.

XX 12-APR-2001.

XX 10-OCT-2000; 2000WO-US028095.

XX 07-OCT-1999; 99US-0158338P.

XX 07-OCT-1999; 99US-0158425P.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Reed S, Houghton RL, Mcneill PD, Dillon DC, Lodes ML;

XX WPI; 2001-290576/30.

XX Vaccinating against Mycobacteria infections in mammals using fusion

XX proteins comprising combinations of heterologous antigens.

XX Example 2; Fig 8; 168pp; English.

XX The sequence encodes Mycobacterium tuberculosis fusion protein, TBRA12-

XX HTCC#1 and includes a His tag at the N-terminus to aid purification.

XX Compositions comprising at least 2 heterologous antigens, as a fusion

XX protein, and vectors expressing the fusion proteins are used as vaccines

XX to prophylactically immunise mammals (especially humans) against

XX infection by Mycobacteria. The compositions contain at least 2

XX heterologous antigens that increase the serological sensitivity of

XX individuals infected with tuberculosis, a disease frequently affecting

XX patients with acquired immunodeficiency disease, AIDS

XX Sequence 1629 BP; 298 A; 500 C; 535 G; 296 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,97e-54 Length: 1629

Score: 675.00 Matches: 132

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 5 Gaps: 0

US-09-684-215B-4 (1-132) x AAS03793 (1-1629)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

DB 25 ACGCCCGCGTCCGATAACTTCCAGCTGTCCAGGTGGGAGGATTCGCCATTCGGATC 84

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

DB 85 GGGCAGCGGATGGCGATTCGCGGCCAGATCCGATCGGGTGGGGTCCACCACCGTTTCA 144

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60

DB 145 ATCGGGCTTACCGCTTCCCTCGGCTTGGGTGTGTCGACAAACGCGACGCGCAGCA 204

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

DB 205 GTCCAAACGGGTGTGCGGAGCGCTCCGCGGCCAGATCTCGGCATCTTCCACCGCGCAGTG 264

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

DB 265 ATCACCGCGTGTGCGGAGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTAA 324

QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleValSerGlyGlyThrArg 120

DB 325 GGGCATCATCCCGTGTGCGGTATCTCGGTGACCTTGGCAACCAAGTGGGCGCGCAGCGT 384

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

```
DB      385 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 420
|||||
RESULT 13
AAL40771
ID      AAL40771 standard; DNA; 1742 BP.
AC      AAL40771;
XX      03-OCT-2002 (first entry)
XX      Nucleotide sequence encoding Ral2-WT1 fusion protein.
DE      Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
XX      vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW      immunogen; cytokine; gene; ds.
XX      Unidentified.
OS      Chimeric.
XX      Key      Location/Qualifiers
FH      CDS      4..1740
FT      /*tag= a
FT      /product= "Ral2-WT1 fusion protein"
XX      WO200125401-A2.
XX      12-APR-2001.
XX      06-OCT-2000; 2000WO-US027652.
XX      07-OCT-1999; 99US-0158585P.
XX      (CORI-) CORIXA CORP.
XX      Skeiky Y, Guderian J;
XX      WPI; 2001-266299/27.
XX      P-PSDB; AAO22140.
XX      Recombinant nucleic acid molecule for producing high yield expression of
FT      desired fusion polypeptides, encodes fusion polypeptide comprising
PT      Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX      Disclosure; Fig 4; 39pp; English.
XX      The invention relates to a recombinant nucleic acid molecule encoding a
CC      fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14
CC      kDa C-terminal fragment of serine protease antigen MTB32A of
CC      Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC      The recombinant fusion nucleic acids and polypeptides are useful for
CC      providing stable and high yield expression of fusion polypeptides of both
CC      eukaryotic and prokaryotic origin and to encode a protein product for use
CC      as an antigen for detecting serum antibodies. The presence of serum
CC      antibodies to M. tuberculosis antigens in an individual indicates that
CC      the individual is infected with it. The fusion polypeptides are useful as
CC      sources of proteins for monitoring binding of serum antibodies to fusion
CC      proteins and as an immunogen to induce and/or enhance immune responses.
CC      The coding sequences can be ligated with a coding sequence of another
CC      molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC      can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC      represents the DNA encoding the Ral2-WT1 fusion protein
XX      Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.:      4,28e-54      Length:      1742
Score:          675.00      Matches:      132
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              4      Gaps:      0
```

```
US-09-684-215B-4 (1-132) x AAL40771 (1-1742)
QY      1 ThrAlaAlaSerAspAsnPhgGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
DB      25 ACGGCCGGTCCGATAAATCTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCGGATC 84
QY      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB      85 GGGCAGCGGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCACCGCTTCAT 144
QY      41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
DB      145 ATCGGGCTACCGCTTCCTCGGCTTGGTGTTCGACAAACGCAACGGCAGCGGCACGA 204
QY      61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB      205 GTCCAAACGGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGGTG 264
QY      81 IleThrAlaValAspGlyValaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB      265 ATCACCAGCGGTCCAGCGGCTCCGATCAACTCGCCACCGCGATGGCGGACGCGCTTAAC 324
QY      101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlySerGlyGlyThrArg 120
DB      325 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGGCAACCAAGTCGGCGGCGACGCT 384
QY      121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB      385 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 420

RESULT 14
AAL40768
ID      AAL40768 standard; DNA; 1871 BP.
XX      AAL40768;
XX      AC
XX      03-OCT-2002 (first entry)
XX      DE      Nucleotide sequence encoding MTB32A protein.
XX      Ral2; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW      vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW      immunogen; cytokine; gene; ds.
XX      Mycobacterium tuberculosis.
XX      Key      Location/Qualifiers
FH      CDS      89..1156
FT      /*tag= a
FT      /product= "Bacillus mycobacterium MTB32A protein"
XX      WO200125401-A2.
XX      12-APR-2001.
XX      06-OCT-2000; 2000WO-US027652.
XX      07-OCT-1999; 99US-0158585P.
XX      (CORI-) CORIXA CORP.
XX      Skeiky Y, Guderian J;
XX      WPI; 2001-266299/27.
XX      P-PSDB; AAO22137.
XX      Recombinant nucleic acid molecule for producing high yield expression of
FT      desired fusion polypeptides, encodes fusion polypeptide comprising
PT      Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX      Disclosure; Fig 1; 39pp; English.
XX      The invention relates to a recombinant nucleic acid molecule encoding a
CC
```

CC fusion polypeptide, comprising a polynucleotide sequence of Ral2, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Mycobacterium tuberculosis MTB32A protein  
 XX  
 XX  
 SQ Sequence 1871 BP; 317 A; 617 C; 605 G; 332 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4,65e-54 Length: 1871  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x AAL40768 (1-1871)

Qy 1 ThrAlaIaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 758 ACGGCCGCGTCCGATTAATCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 817  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 Db 818 GGGCAGCGCATGGCGATCGGGGGCCAAATCCGATCGGGTGGGGGTCCACCCCGTTCAT 877  
 Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlnGlyValAlaArg 60  
 Db 878 ATCCGGCTACCGGCTTCTCCGGTTGGTGTTCGACACACGCGACGGCGACGA 937  
 Qy 61 ValGlnArgValAlGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 938 GTCCAAACGCGTGGTCGGAAGCGTCCGCGGCGCAAGTCTCGCATCTCCACCGCGCAGGTG 997  
 Qy 81 IleThrAlaValAlaArgGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 998 ATCCCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGGCCTTAAC 1057  
 Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 1058 GGGCATCATCCCGGTGACGTCTATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 1117  
 Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 1118 ACAGGGAACGTGACATTCGGCGAGGACCCCGGCC 1153

RESULT 15  
 ADA26353  
 ID ADA26353 standard; DNA; 2181 BP.  
 AC  
 XX ADA26353;  
 AC  
 XX 20-NOV-2003 (first entry)  
 DT  
 XX Mycobacterium MTB32-MTB39F fusion protein encoding DNA.  
 DE  
 XX ds; gene; fusion polypeptide; MTB32A; MTB39; antigen; MTB32A; MTB39;  
 KW MTB35A; tuberculosis; tuberculostatic; gene therapy; vaccine.  
 XX  
 XX Chimeric.  
 OS Mycobacterium sp.  
 XX  
 XX Key Location/Qualifiers

CDS 4. .2175  
 /\*tag= a  
 /product= "MTB32-MTB39F fusion polypeptide"

WO2003070187-A2.

28-AUG-2003.

18-FEB-2003; 2003WO-US004903.

15-FEB-2002; 2002US-0357351P.

(CORI-) CORIXA CORP.

Skeiky Y, Guderian J, Reed S;

WPI: 2003-697554/66.

P-PSDB; ADA26354.

XX New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
 PT and MTB39 antigens, with or without the MTB85A antigen, from a  
 PT Mycobacterium species, useful for diagnosing, preventing and/or treating  
 PT tuberculosis infection.

PS Claim 1; Fig 1; 112pp; English.

XX The invention relates to a novel nucleic acid encoding a fusion  
 CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
 CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
 CC A polypeptide of the invention has tuberculostatic activity. A  
 CC polynucleotide of the invention may have a use in gene therapy, and as a  
 CC vaccine. The methods and compositions of the present invention are useful  
 CC for diagnosing, preventing and/or treating tuberculosis infection. The  
 CC present sequence is used in the exemplification of the invention.

XX Sequence 2181 BP; 341 A; 693 C; 792 G; 355 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.53e-54 Length: 2181  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-09-684-215B-4 (1-132) x ADA26353 (1-2181)

Qy 1 ThrAlaIaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 598 ACGGCCGCGTCCGATTAATCTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 657  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 Db 658 GGGCAGCGCATGGCGATCGGGGGCCAGATCCGATCGGGTGGGGGTCCACCCCGTTCAT 717  
 Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlnGlyValAlaArg 60  
 Db 718 ATCCGGCTACCGGCTTCTCCGGTTGGTGTTCGACACACGCGACGGCGACGA 777  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 778 GTCCAAACGCGTGGTCGGAAGCGTCCGCGGCGCAAGTCTCGCATCTCCACCGCGCAGGTG 837  
 Qy 81 IleThrAlaValAlaArgGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 838 ATCCCGCGGTGCGACGGCGCTCCGATCAACTCGGCCACCGGATGGCGGACGGCCTTAAC 897  
 Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 898 GGGCATCATCCCGGTGACGTCTATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 957  
 Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

958 ACAGGGAACTGACATGGCCGAGGAGCCCGCC 993

RESULT 16

AD47084  
ID AD47084 standard; DNA; 2190 BP.  
XX  
AC AD47084;  
XX  
29-AUG-2003 (revised)  
DT 27-JAN-2003 (first entry)  
XX  
XX Mycobacterium sp. MTB72FmutSA fusion protein encoding DNA.  
XX  
XX vaccine; immunity; diagnostic agent; gene therapy; TBH9; antigen;  
KW Ra35MutSA; Ra12; MTB72MutSA; chimeric; gene; ds.  
XX  
XX Mycobacterium sp.  
OS Mycobacterium tuberculosis.  
OS Chimeric.

XX Key Location/Qualifiers  
FH 1. .2190  
FT CDS /tag= a  
FT /product= "MTB72FmutSA fusion protein"

WO200272792-A2.

19-SEP-2002.

13-MAR-2002; 2002WO-US008223.

13-MAR-2001; 2001US-0275837P.

(CORI-) CORIXA CORP.

Skelky Y, Brannon M, Cuderian J;

WPI; 2002-759844/82.

P-PSDB; AAE29709.

New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity against pathogenic microorganisms e.g. Leishmania and Mycobacterium tuberculosis.

Disclosure; Page 92-93; 155pp; English.

The invention relates to a recombinant nucleic acid molecule encoding a fusion polypeptide. The recombinant nucleic acid comprises a heterologous polynucleotide sequence encoding an antigen or an antigenic fragment from Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a polypeptide or its fragment. The Leishmania polynucleotide is selected from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention are used in methods for eliciting immune response in mammals. They are useful as vaccines to elicit protective immunity against pathogenic microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion polypeptides are used for enhancing the expression of polynucleotides, as in vivo diagnostic agents and for raising antibodies in a non-human animal. The invention is used in gene therapy. The present sequence is a DNA encoding MTB72F fusion protein. This fusion protein comprises Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and TBH9 protein from Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,55e-54 Length: 2190  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps:

US-09-684-215B-4 (1-132) x AAD47084 (1-2190)

QY 1 ThrAlaLaSerAspAenPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
DB 22 ACGCCCGGTCGGATACCTCCAGCTGTCCAGGGTGGCAGGATTCCCAATCCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
DB 82 GGGCAGGGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCAACCCCGTTTCAT 141  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAenAsnGlyAsnGlyAlaArg 60  
DB 142 ATCGGGCTACCGCTTCTCGGCTTGGGTGTTCGACAAACAACGGCAACGGCGCAGA 201  
QY 61 ValGlnArgValValGlySerAlaAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
DB 202 GTCCAAACGGTGTGGAGGGCTCCGGGGGAGTCTCGGCATCTCCACCGGCGACGTG 261  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
DB 262 ATCACCGGGTCCAGCGGGCTCCGATCAACTCGGCCACCGCGATGGCGACGCGTTAAC 321  
QY 101 GlyHisHisProGlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArg 120  
DB 322 GGGCATCATCCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCGGGGCGACGGGT 381

RESULT 17

AAD28343  
ID AAD28343 standard; DNA; 2190 BP.

XX AAD28343;

22-APR-2002 (first entry)

Mycobacterium species MTB72FmutSA fusion protein encoding DNA.

Fusion protein; antigen; serological sensitivity; immune response;  
tuberculosis; infection; vaccine; MTB72FmutSA; Ra12-TBH9-Ra35MutSA;  
mutant; mutein; ds.

Mycobacterium sp.

XX Key Location/Qualifiers  
FH 1. .2190  
FT CDS /tag= a  
FT /product= "MTB72FmutSA fusion protein"  
FT misc\_feature /tag= b  
FT /note= "Ra12 DNA fragment"  
FT misc\_feature /tag= c  
FT /note= "TBH9FL DNA fragment"  
FT misc\_feature /tag= d  
FT /note= "Ra35 DNA fragment"  
FT mutation /tag= e  
FT replace(2128, T)

WO200198450-A2.

27-DEC-2001.

20-JUN-2001; 2001WO-US019959.

20-JUN-2000; 2000US-00597796.

01-FEB-2001; 2001US-0265737P.

(CORI-) CORIXA CORP.

XX FI Skeiky Y, Reed S, Alderson M;  
 XX WPI; 2002-147798/19.  
 DR DR P-PSDB; AAEL17573.  
 XX  
 PT Composition comprising MTB39 antigen and MTB32A antigen from  
 PT Mycobacterium species, useful for eliciting immune response in a subject.  
 XX  
 PS Claim 81; Page 108-109; 136pp; English.  
 XX  
 CC The present invention relates to fusion proteins containing at least two  
 CC Mycobacterium species antigens, nucleotides encoding them and  
 CC compositions comprising such fusion proteins. The present invention  
 CC particularly relates to nucleic acids encoding fusion proteins that  
 CC include two or more individual M. tuberculosis antigens which increase  
 CC the serological sensitivity of sera from individuals infected with  
 CC tuberculosis and methods for their use in diagnosis, prevention and  
 CC treatment of tuberculosis infection. Sequences of the invention are  
 CC useful for eliciting an immune response in a mammal, e.g., human,  
 CC immunised with BCG. They are useful in the diagnosis, treatment and  
 CC prevention of Mycobacterium infection. The fusion proteins and the  
 CC polynucleotides are useful as diagnostic tools in patients infected with  
 CC Mycobacterium, in vitro and in vivo assays for detecting humoral  
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the  
 CC diagnosis of an infection or monitoring of disease progression, as  
 CC immunogens to generate or elicit a protective immune response in a  
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human  
 CC animal. Sequences of the invention are also used as vaccines. MTB32A  
 CC fusion proteins of the invention are useful as in vivo diagnostic agents  
 CC for intradermal skin test. The present sequence is a DNA encoding  
 CC Mycobacterium species MTB72FMutSA (Ra12-TD9-Ra3MutSA) mutant fusion  
 CC protein  
 XX  
 SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5,56e-54 Length: 2190  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 DB:  
 US-09-684-215B-4 (1-132) x AAD28343 (1-2190)  
 QY 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 22 ACGGCGCGTCCGATAACTTCAGCTGTCCAGGCTGGCGAGGATTCGCGATTCGCGATC 81  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 82 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTGGGGGTCCACCGCTTCAT 141  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArg 60  
 Db 142 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTGTGCGACAAACACGCGACGGCGACGA 201  
 QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 202 GTTCAACGCGTGTGGGAGGCGCTCCGCGGCAAGTCTCGGCAATTCACCGGCGCGTG 261  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 262 ATCACCGGCTGACGCGGCTCCGATCACTCGGCGACCGGATGGCGAGCGCTTAAC 321  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 322 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGCGGCGCGGT 381  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 382 ACAGGAGACGTGACATTGGCCGAGGAGCCCCCGGCC 417

RESULT 18  
 AAL40773  
 ID AAL40773 standard; DNA; 2191 BP.  
 XX  
 AC AAL40773;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Nucleotide sequence encoding Ra12-H9-32A fusion protein.  
 XX  
 KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;  
 KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;  
 KW immunogen; cytokine; gene; ds.  
 XX  
 OS Unidentified.  
 OS Chimeric.  
 FH Key Location/Qualifiers  
 CDS 1..2190  
 FT /\*tag= a  
 FT /product= "Ra12-H9-32A fusion protein"  
 FT  
 XX WO200125401-A2.  
 PN 12-APR-2001.  
 PD 06-OCT-2000; 2000WO-US027652.  
 PF 07-OCT-1999; 99US-0158585P.  
 PR (CORI-) CORIXA CORP.  
 PA  
 XX Skeiky Y, Guderian J;  
 XX WPI; 2001-266299/27.  
 DR P-PSDB; AAO22142.  
 DR  
 XX Recombinant nucleic acid molecule for producing high yield expression of  
 PT desired fusion polypeptides, encodes fusion polypeptide comprising  
 PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.  
 XX  
 XX Disclosure; Fig 6; 39pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14  
 CC kDa C-terminal fragment of serine protease antigen MTB32A of  
 CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.  
 CC The recombinant fusion nucleic acids and polypeptides are useful for  
 CC providing stable and high yield expression of fusion polypeptides of both  
 CC eukaryotic and prokaryotic origin and to encode a protein product for use  
 CC as an antigen for detecting serum antibodies. The presence of serum  
 CC antibodies to M. tuberculosis antigens in an individual indicates that  
 CC the individual is infected with it. The fusion polypeptides are useful as  
 CC sources of proteins for monitoring binding of serum antibodies to fusion  
 CC proteins and as an immunogen to induce and/or enhance immune responses.  
 CC The coding sequences can be ligated with a coding sequence of another  
 CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and  
 CC can be used in vivo as a DNA vaccine. This polynucleotide sequence  
 CC represents the DNA encoding the Ra12-H9-32A fusion protein  
 XX  
 SQ Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 1 Other;  
 Alignment Scores:  
 Pred. No.: 5,56e-54 Length: 2191  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 4  
 DB:  
 US-09-684-215B-4 (1-132) x AAL40773 (1-2191)





XX Tuberculosis; antigen; fusion protein; Mtb32A; Ra12; TBH9; Ra35;  
 KW diagnosis; therapy; vaccine; immunogen; ss.

XX Mycobacterium tuberculosis.

XX Key Location/Qualifiers  
 FT CDS 42..2231  
 FT /\*tag= a

XX W09951748-A2.

XX 14-OCT-1999.

XX 07-APR-1999; 99WO-US007717.

XX 07-APR-1998; 98US-00056556.

XX 30-DEC-1998; 98US-00223040.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Alderson M, Campos-Neto A;

XX WPI; 1999-601610/51.

XX P-PSDB; AAY32059.

XX New fusion proteins useful for diagnosis, prevention and treatment of  
 FT tuberculosis.

XX Example; Fig 1A-B; 83pp; English.

XX This DNA sequence includes a coding region for a recombinant  
 CC Mycobacterium tuberculosis tri-antigen fusion protein (see AAY32059),  
 CC termed Mtb32A, composed of the antigens Ra12, TBH9 and Ra35. The DNA is  
 CC useful for the recombinant production of the fusion protein. Coding  
 CC sequences for the antigens were modified by PCR in order to facilitate  
 CC their fusion and subsequent expression of the fusion protein. 3 Coding  
 CC sequences for Ra12, TBH9 and Ra25 were ligated to encode Mtb32A. The  
 CC invention provides fusion proteins (see AAY32059-71) containing at least  
 CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides  
 CC encoding them are useful as vaccines for preventing tuberculosis  
 CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests  
 CC for detection of anti-M. tuberculosis antibodies), monitoring of disease  
 CC progression, and treatment of tuberculosis. They are more effective  
 CC immunogens than mixtures of the individual protein components

XX SQ Sequence 2287 BP; 374 A; 723 C; 805 G; 382 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 5.83e-54 Length: 2287  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-684-215B-4 (1-132) x AAZ20194 (1-2287)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 63 ACGCGCGTCCGATACCTTCCAGTCTCCAGGTCGGCAGGATTCGCATCCGATC 122  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 DB 123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGTCGGGGGTCCACCCCGTTTCAT 182  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAsnAnGlyValAlaArg 60  
 DB 193 ATCGGGCTACCGCTTCTCGCTTGGTGTGTTGTCGACACACACGCGCAGCGACGA 242  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 243 GTCCAAACGGTGTGGGAGGCGCTCCGCGGCAAGTCTCGGCATCTCCACCGCGCGCGTG 302

QY 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 303 ATCACCAGGCTCGACGGGCTCCGATCACTCGCCACCGGATGGGACGCGCTTAAC 362  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 363 GGGCATCATCCCGGTGACGTCTCTCGTGACCTCGGAAACCAAGTCGGCGGCGACGCGT 422  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 423 ACAGGGAACGTGACATTCGCCGAGGAGGCCCGCGGCC 458

RESULT 21

AAD47083

ID AAD47083 standard; DNA; 2287 BP.

XX AAD47083;

XX 29-AUG-2003 (revised)

DT 27-JAN-2003 (first entry)

XX Mycobacterium sp. MTB72F fusion protein encoding DNA.

XX Vaccine; immunity; diagnostic agent; gene therapy; TBH9; antigen; Ra35;  
 KW Ra12; MTB72F; chimeric; gene; ds.

XX Mycobacterium sp.

OS Mycobacterium tuberculosis.

OS Chimeric.

XX Key Location/Qualifiers

FT CDS 42..2231

FT /\*tag= a

FT /product= "MTB72F fusion protein"

XX W0200272792-A2.

XX 19-SEP-2002.

XX 13-MAR-2002; 2002WO-US008223.

XX 13-MAR-2001; 2001US-0275837P.

XX (CORI-) CORIXA CORP.

XX Skeiky Y, Brannon M, Guderian J;

XX WPI; 2002-759844/82.

XX P-PSDB; AAE29708.

XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,  
 PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity  
 PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium  
 PT tuberculosis.

XX Disclosure; Page 87-90; 155pp; English.

XX The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
 CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
 CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
 CC are used in methods for eliciting immune response in mammals. They are  
 CC useful as vaccines to elicit protective immunity against pathogenic  
 CC microorganisms such as Leishmania and Mycobacterium tuberculosis, as  
 CC polypeptides are used for enhancing the expression of polynucleotides, as  
 CC in vivo diagnostic agents and for raising antibodies in a non-human  
 CC animal. The invention is used in gene therapy. The present sequence is a  
 CC DNA encoding MTB72F fusion protein. This fusion protein comprises Ra12  
 CC and Ra35 protein from Mycobacterium tuberculosis and TBH9 protein from  
 CC Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS field)

```
XX SQ Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;
Alignment Scores:
Pred. No.: 5,83e-54 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 6
US-09-684-215B-4 (1-132) x AAD47083 (1-2287)
Qy 1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 63 ACGCCCGCTCCGATACCTCCAGCTGTCAGGGTGGGAGGATCGCATTCGGATC 122
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 123 GGGCAGGCGATGCGGATCGCGGCGAGATCGATCGGGTGGGGGTCCACCCAGCTTCAT 182
Qy 41 IleGlyProThrAlaPheLeuGlyGlnIleArgSerGlyGlySerProThrValHis 60
Db 183 ATCCAGCGCTACCGCTTCTCCGCTTGGGTGTTGTCACCAACCGGCAACGGCGACGA 242
Qy 61 ValGlnArgValValGlySerAlaProAlaIleSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCAAACGCGTGGTCCGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGGCGAGTG 302
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCAGCGTGGTCCGAGCGCTCCGATCACTCGGCCACCGGATGGCGGACGCTTAAC 362
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120
Db 363 GGGCATCATCCGCTGACGTATCTCGGTGACCTGGCAACCAAGTGGGCGGCGACGCT 422
Qy 121 ThrGlyAsnValThrLeuAlaGlyGlnProAla 132
Db 423 ACAGGACGTCGATTCGCGGAGGACCCCGGCC 458
RESULT 22
RAD28342
ID AAD28342 standard; DNA; 2287 BP.
XX AC AAD28342;
XX DT 22-APR-2002 (first entry)
XX DE Mycobacterium species MTB72F fusion protein encoding DNA.
XX KW Fusion protein; antigen; serological sensitivity; immune response;
XX OS Mycobacterium sp.
XX FH Location/Qualifiers
XX FT 42..2231
XX FT CDS /tag= a
XX FT /product= "MTB72F fusion protein"
XX FT misc_feature /tag= b
XX FT /note= "Ra12 DNA fragment"
XX FT misc_feature /tag= c
XX FT /note= "TbH9FL DNA fragment"
XX FT misc_feature /tag= d
XX FT /note= "Ra35 DNA fragment"
XX PN WO200198460-A2.
XX PD 27-DEC-2001.
```

```
XX PF 20-JUN-2001; 2001WO-US019959.
XX PR 20-JUN-2000; 2000US-00597796.
XX PR 01-FEB-2001; 2001US-0265737P.
XX PA (CORI-) CORIXA CORP.
XX SK Skeiky Y, Reed S, Alderson M;
XX WI 2002-147798/19.
XX DR P-PSDB; AAB17572.
XX Composition comprising MTB39 antigen and MTB32A antigen from
XX Mycobacterium species, useful for eliciting immune response in a subject.
XX Claim 62; Page 103-106; 136pp; English.
XX The present invention relates to fusion proteins containing at least two
XX Mycobacterium species antigens, nucleotides encoding them and
XX compositions comprising such fusion proteins. The present invention
XX particularly relates to nucleic acids encoding fusion proteins that
XX include two or more individual M. tuberculosis antigens which increase
XX the serological sensitivity of sera from individuals infected with
XX tuberculosis and methods for their use in diagnosis, prevention and
XX treatment of tuberculosis infection. Sequences of the invention are
XX useful for eliciting an immune response in a mammal, e.g., human,
XX immunised with BCG. They are useful in the diagnosis, treatment and
XX prevention of Mycobacterium infection. The fusion proteins and the
XX polynucleotides are useful as diagnostic tools in patients infected with
XX Mycobacterium, in vitro and in vivo assays for detecting humoral
XX antibodies or cell-mediated immunity against M. tuberculosis, for the
XX diagnosis of an infection or monitoring of disease progression, as
XX immunogens to generate or elicit a protective immune response in a
XX patient and for raising anti-M. tuberculosis antibodies in a non-human
XX animal. Sequences of the invention are also used as vaccines. MTB32A
XX fusion proteins of the invention are useful as in vivo diagnostic agents
XX for intradermal skin test. The present sequence is a DNA encoding
XX Mycobacterium species MTB72F (Ra12-TbH9-Ra35) fusion protein
XX SQ Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;
Alignment Scores:
Pred. No.: 5,83e-54 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 6
US-09-684-215B-4 (1-132) x AAD28342 (1-2287)
Qy 1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 63 ACGCCCGCTCCGATACCTCCAGCTGTCAGGGTGGGAGGATCGCATTCGGATC 122
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 123 GGGCAGGCGATGCGGATCGCGGCGAGATCGATCGGGTGGGGGTCCACCCAGCTTCAT 182
Qy 41 IleGlyProThrAlaPheLeuGlyGlnIleArgSerGlyGlySerProThrValHis 60
Db 183 ATCCAGCGCTACCGCTTCTCCGCTTGGGTGTTGTCACCAACCGGCAACGGCGACGA 242
Qy 61 ValGlnArgValValGlySerAlaProAlaIleSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCAAACGCGTGGTCCGAGCGCTCCGATCACTCGGCCACCGGATGGCGGACGCTTAAC 302
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCAGCGTGGTCCGAGCGCTCCGATCACTCGGCCACCGGATGGCGGACGCTTAAC 362
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120
```

Db 363 GGGCATCATCCCGTGAAGTCATCTCGTGAGCTGGCAACCAAGTCGGGGCGCACGGT 422  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 423 ACAGGGAACGTGACATTCGGCCGAGGACCCCGGCC 458  
RESULT 23  
ADA26360  
ID ADA26360 standard; DNA; 2451 BP.  
XX  
AC ADA26360;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Mycobacterium MTB72F-DPV (fusion MTB81F) protein encoding DNA.  
XX  
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
KW tuberculosis; tuberculostatic; gene therapy; vaccine.  
XX  
OS Chimeric.  
OS Mycobacterium sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 4..2445  
FT /tag= a  
FT /product= " MTB72F-DPV (fusion MTB81F) protein"  
XX  
PN WO2003070187-A2.  
XX  
PD 28-AUG-2003.  
XX  
PF 18-FEB-2003; 2003WO-US004903.  
XX  
PR 15-FEB-2002; 2002US-0357351P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky Y, Guderian J, Reed S;  
XX  
DR WPI; 2003-697554/66.  
DR P-PSDB; ADA26367.  
XX  
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
PT and MTB39 antigens, with or without the MTB85A antigen, from a  
PT Mycobacterium species, useful for diagnosing, preventing and/or treating  
PT tuberculosis infection.  
XX  
PS Claim 84; Fig 8; 112pp; English.  
XX  
CC The invention relates to a novel nucleic acid encoding a fusion  
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
CC A polypeptide of the invention has tuberculostatic activity. A  
CC vaccine. The methods and compositions of the present invention are useful  
CC for diagnosing, preventing and/or treating tuberculosis infection. The  
CC present sequence is used in the exemplification of the invention.  
XX  
SQ Sequence 2451 BP; 389 A; 785 C; 870 G; 407 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 6.31e-54 Length: 2451  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
  
US-09-684-215B-4 (1-132) x ADA26360 (1-2451)  
QY 1 ThrAlaAlaSerAspAsnThrGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 25 ACGCCCGCTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCCGCAATTCGGATC 84  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGGCGATGGCAATCGCGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 144  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAenGlyAenGlyAlaArg 60  
Db 145 ATCGGGCTACCGCTTCTCGGCTTGGTGTGTGTCACAAACACGCAACGGCGCACGA 204  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCAACGCGTGGTGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACGGCGACGTG 264  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACC CGCGTCGACGGCGCTCCGATCACTCGGCCACCGCATGGCGGACGCGCTTAAC 324  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyvsSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTCTCGGTGACCTGGCAACCAAGTCGGGGCGGACCGCT 384  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTCGGCGGAGGACCCCGGCC 420  
RESULT 24  
ADA26359  
ID ADA26359 standard; DNA; 2487 BP.  
XX  
AC ADA26359;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Mycobacterium MTB72F-MTI (fusion MTB83F) protein encoding DNA.  
XX  
KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
KW tuberculosis; tuberculostatic; gene therapy; vaccine.  
XX  
OS Chimeric.  
OS Mycobacterium sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 4..2481  
FT /tag= a  
FT /product= "MTB72F-MTI (fusion MTB83F) protein"  
XX  
PN WO2003070187-A2.  
XX  
PD 28-AUG-2003.  
XX  
PF 18-FEB-2003; 2003WO-US004903.  
XX  
PR 15-FEB-2002; 2002US-0357351P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky Y, Guderian J, Reed S;  
XX  
DR WPI; 2003-697554/66.  
DR P-PSDB; ADA26366.  
XX  
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
PT and MTB39 antigens, with or without the MTB85A antigen, from a  
PT Mycobacterium species, useful for diagnosing, preventing and/or treating  
PT tuberculosis infection.  
XX  
PS Claim 84; Fig 7; 112pp; English.  
XX  
CC The invention relates to a novel nucleic acid encoding a fusion  
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
CC A polypeptide of the invention has tuberculostatic activity. A



DE Mycobacterium sp. MTB72F-Leishmania sp. MAPS (aka r95f)-fusion DNA.  
 XX Vaccine; immunity; diagnostic agent; gene therapy; MTB72F; MAPS;  
 KW chimeric; gene; ds.  
 XX  
 OS Mycobacterium sp.  
 OS Leishmania sp.  
 OS Chimeric.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 4..2796  
 FT /\*tag= a  
 FT /product= "Mycobacterium sp. MTB72F-Leishmania sp. MAPS  
 FT (aka r95f) fusion protein"  
 XX  
 PN WO200272792-A2.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 13-MAR-2002; 2002WO-US008223.  
 XX  
 PR 13-MAR-2001; 2001US-0275837P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skeiky Y, Brannon M, Guderian J;  
 XX  
 DR WPI; 2002-759844/82.  
 DR P-PSDB; AAE29731.  
 XX  
 XX New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,  
 PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity  
 PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium  
 PT tuberculosis.  
 XX  
 XX Example 6; Page 128-129; 155pp; English.  
 XX  
 CC The invention relates to a recombinant nucleic acid molecule encoding a  
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous  
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from  
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a  
 CC polypeptide or its fragment. The Leishmania polynucleotide is selected  
 CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention  
 CC are used in methods for eliciting immune response in mammals. They are  
 CC used as vaccines to elicit protective immunity against pathogenic  
 CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion  
 CC polypeptides are used for enhancing the expression of polynucleotides, as  
 CC in vivo diagnostic agents and for raising antibodies in a non-human  
 CC animal. The invention is used in gene therapy. The present sequence is  
 CC Mycobacterium sp. MTB72F-Leishmania sp. thiol specific antioxidant [TSA;  
 CC MAPS (aka r95f)] fusion DNA. This sequence comprises Mycobacterium sp.  
 CC MTB72F (a 72 kDa poly-protein fusion construct comprising Ral2-TbH9-Ra35)  
 CC linked to the Leishmania TSA (MAPS). (Updated on 29-AUG-2003 to  
 CC standardise OS field)  
 XX  
 XX Sequence 2808 BP; 465 A; 889 C; 982 G; 471 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 7.36e-54 Length: 2808  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-684-215B-4 (1-132) x AAD47110 (1-2808)  
 QY 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
 DB 25 ACGGGCGGTCCGATTAACCTTCAGGTGTCACAGGGTGGCAGGATTCGCATTCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

DB 85 GGGCAGCGGATCGGATCGCGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTTCAT 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValaArg 60  
 DB 145 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTGTGACAAACCGGCAACGGCGACGA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCACACGGCTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGACGTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 265 ATCACCAGGTCGACGGCGCTCCGATCACTCGGCCACCGCATGGCGGACGGCCTTAAC 324  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTriPnThrLyssSerGlyGlyThrArg 120  
 DB 325 GGGCATCATCCGGTGACGTCACTCGTGACCTGGCAACCAAGTCGGCGGCGACGCGT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGGAACGTGACATTGGCGGAGGAGGACCCCGGCC 420  
 RESULT 27  
 ADA26357  
 ID ADA26357 standard; DNA; 2808 BP.  
 XX  
 AC ADA26357;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Mycobacterium MTB72F-MAPS (fusion r95f) protein encoding DNA.  
 XX  
 KW ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
 KW tuberculosis; tuberculostatic; gene therapy; vaccine.  
 XX  
 OS Chimeric.  
 OS Mycobacterium sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 4..2796  
 FT /\*tag= a  
 FT /product= "MTB72F-MAPS (fusion r95f)"  
 XX  
 PN WO2003070187-A2.  
 XX  
 PD 28-AUG-2003.  
 XX  
 PF 18-FEB-2003; 2003WO-US004903.  
 XX  
 PR 15-FEB-2002; 2002US-0357351P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skeiky Y, Guderian J, Reed S;  
 XX  
 DR WPI; 2003-697554/66.  
 DR P-PSDB; ADA26364.  
 XX  
 XX New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
 PT and MTB39 antigens, with or without the MTB85A antigen, from a  
 PT Mycobacterium species, useful for diagnosing, preventing and/or treating  
 PT tuberculosis infection.  
 XX  
 PS Disclosure; Fig 5; 112pp; English.  
 XX  
 CC The invention relates to a novel nucleic acid encoding a fusion  
 CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
 CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
 CC A polypeptide of the invention has tuberculostatic activity. A  
 CC polynucleotide of the invention may have a use in gene therapy, and as a  
 CC vaccine. The methods and compositions of the present invention are useful  
 CC for diagnosing, preventing and/or treating tuberculosis infection. The  
 CC present sequence is used in the exemplification of the invention.

XX SQ Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.36e-54 Length: 2808  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-09-684-215B-4 (1-132) x ADA26357 (1-2808)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 25 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGCATTCGGATC 84  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGGATGGCGATCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 144  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
Db 145 ATCGGGCTACCGGCTTCCTCGGCTTGGTGTTCGACAAACACGGAACGGCGACGA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCCAAACGCGTGGTGGGAGCGTCCGCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 264  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACCAGCGTGGAGCGGCTCCGATCACTCGGCGCACCGGATGGCGGACGGCTTAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGTGGTGCATCTCTCGTGCACCTGGCAACCAAGTCCGGCGGCACGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 420

RESULT 28

ADA26355

ID ADA26355 standard; DNA; 3030 BP.

XX AC ADA26355;

XX DT 20-NOV-2003 (first entry)

XX Mycobacterium MTB-102F fusion protein encoding DNA.

XX ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
XX tuberculosis; tuberculostatic; gene therapy; vaccine.

XX Chimeric.

XX Mycobacterium sp.

XX Key Location/Qualifiers  
FT 1..3030

FT /tag= a  
FT /product= "MTB-102F fusion protein"  
FT /note= "No stop codon given"

XX WO2003070187-A2.

XX PD 28-AUG-2003.

XX PF 18-FEB-2003; 2003WO-US0004903.

XX PR 15-FEB-2002; 2002US-0357351P.

XX PA (CORI-) CORIXA CORP.

XX

PI Skeiky Y, Guderian J, Read S;

XX WPI; 2003-697554/66.  
XX P-PSDB; ADA26356.

XX New isolated nucleic acid encoding a fusion polypeptide with the MTB32A  
PT and MTB39 antigens, with or without the MTB85A antigen, from a  
PT Mycobacterium species, useful for diagnosing, preventing and/or treating  
PT tuberculosis infection.

XX Claim 5; Fig 3; 112pp; English.

XX The invention relates to a novel nucleic acid encoding a fusion  
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and  
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.  
CC A polypeptide of the invention has tuberculostatic activity. A  
CC polynucleotide of the invention may have a use in gene therapy, and as a  
CC vaccine. The methods and compositions of the present invention are useful  
CC for diagnosing, preventing and/or treating tuberculosis infection. The  
CC present sequence is used in the exemplification of the invention.

XX SQ Sequence 3030 BP; 498 A; 966 C; 1062 G; 504 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.03e-54 Length: 3030  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-09-684-215B-4 (1-132) x ADA26355 (1-3030)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 595 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGGAGGATTCGCCATTCGGATC 654  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 655 GGGCAGCGGATGGCGATCCGCGGCGCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 714  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
Db 715 ATCGGCGCTACCGGCTTCCTCGGCTTGGTGTTCGACAAACACGGAACGGCGACGA 774  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 775 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGGCGAGTG 834  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 835 ATCACCAGCGTCCGACGGCGTCCGATCAACTCGGCGCACCGGATGGCGGACGGCTTAAC 894  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 895 GGGCATCATCCCGTGGTGCATCTCGTGCACCTGGCAACCAAGTCCGGCGGCACGCGT 954  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 955 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 990

RESULT 29

ADA26363

ID ADA26363 standard; DNA; 3060 BP.

XX AC ADA26363;

XX DT 20-NOV-2003 (first entry)

XX M. bovis MTB72F and 85b complex antigen (fusion MTB103F) DNA.

XX ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;  
XX tuberculosis; tuberculostatic; gene therapy; vaccine; MTB72F; 85b.

KW

```
XX OS Chimeric.
XX OS Mycobacterium bovis.
XX PH Location/Qualifiers
XX Key 4..3054
XX CDS /*tag= a
XX FT /product= "MTB72F and 85b complex (fusion MTB103P)"
XX PN WO2003070187-A2.
XX PD 28-AUG-2003.
XX PF 18-FEB-2003; 2003WO-US004903.
XX PR 15-FEB-2002; 2002US-0357351P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky Y, Guderian J, Reed S;
XX WPI; 2003-697554/66.
XX DR P-PSDB; ADA26370.
XX FT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
XX PT and MTB39 antigens, with or without the MTB85A antigen, from a
XX PT Mycobacterium species, useful for diagnosing, preventing and/or treating
XX PT tuberculosis infection.
XX PS Claim 84; Fig 11; 112pp; English.
XX CC The invention relates to a novel nucleic acid encoding a fusion
XX CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
XX CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
XX CC A polypeptide of the invention has tuberculostatic activity. A
XX CC polynucleotide of the invention may have a use in gene therapy, and as a
XX CC vaccine. The methods and compositions of the present invention are useful
XX CC for diagnosing, preventing and/or treating tuberculosis infection. The
XX CC present sequence is used in the exemplification of the invention.
XX SQ Sequence 3060 BP; 503 A; 977 C; 1070 G; 510 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.12e-54 Length: 3060
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-684-215B-4 (1-132) x ADA26363 (1-3060)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 25 ACGGCGGTCGATACATTCAGTGTCCAGGGTGGCAGGGATTCGCCATTCGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 85 GGGCAGGCGATGCGCATCGCGGCCAGATCGATCGGTGGGGGTCACCCACCGTTTCAT 144
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60
DB 145 ATCGGGCTACCGCTTCCTGGCTGGTGTTCGACACACACGCGACGGCGCAGA 204
QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 205 GTCCACACGCGTGTGGGAGCGCTCCGCGCGGCAAGTCTCGGCATCTCCACCGGCGCGTG 264
QY 81 IleThrAlaValAlaGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 265 ATCACCGGTCGACGCGCTCCGATCACTCGGCACCGCATCGCGACGCGCGCTTAC 324
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpInThrLysSerGlyGlyThrArg 120
```

```
DB 325 GGGCATCATCCGGTGACGTCTCGGTGACCTGGGAAACAAAGTCGGCGGCACGCGT 384
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGCGCC 420

RESULT 30
ADA26362
ID ADA26362 standard; DNA; 3104 BP.
XX ADA26362;
XX 20-NOV-2003 (first entry)
XX Mycobacterium MTB72F-hTCC#1 (fusion MTB102tm2F) protein encoding DNA.
XX ds; Gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
XX tuberculosis; tuberculostatic; gene therapy; vaccine.
XX Chimeric.
XX OS Mycobacterium sp.
XX Key Location/Qualifiers
XX CDS 4..3072
XX FT /*tag= a
XX FT /product= "MTB72F-hTCC#1 (fusion MTB102tm2F) protein"
XX PN WO2003070187-A2.
XX PD 28-AUG-2003.
XX PF 18-FEB-2003; 2003WO-US004903.
XX PR 15-FEB-2002; 2002US-0357351P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky Y, Guderian J, Reed S;
XX WPI; 2003-697554/66.
XX DR P-PSDB; ADA26369.
XX FT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
XX PT and MTB39 antigens, with or without the MTB85A antigen, from a
XX PT Mycobacterium species, useful for diagnosing, preventing and/or treating
XX PT tuberculosis infection.
XX PS Claim 84; Fig 10; 112pp; English.
XX CC The invention relates to a novel nucleic acid encoding a fusion
XX CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
XX CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
XX CC A polypeptide of the invention has tuberculostatic activity. A
XX CC polynucleotide of the invention may have a use in gene therapy, and as a
XX CC vaccine. The methods and compositions of the present invention are useful
XX CC for diagnosing, preventing and/or treating tuberculosis infection. The
XX CC present sequence is used in the exemplification of the invention.
XX SQ Sequence 3104 BP; 510 A; 968 C; 1104 G; 522 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.25e-54 Length: 3104
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-684-215B-4 (1-132) x ADA26362 (1-3104)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
```



Db 25 ACGGCCGCGTCCGATACCTCCAGCTGTCCAGGGTGGCAGGGATTCCGCATTCGGATC 84  
Qy 21 GlyClnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGCATGGGATCGCGGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTCT 144  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 145 ATCGGGCTTACCGGCTTCTCGGCTTGGGTGTGTCACAAACACGCAACGGCGCACCGA 204  
Qy 61 ValClnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCCACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGACGTG 264  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACC GCGGTCGACGCGCTCCGATCACTCGGCCACCGCATGGCGACGGCGCTTAA 324  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIlysserGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGGCAAAACCAAGTCGGCGGCGCGGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC 420

Search completed: April 29, 2004, 23:03:02  
Job time : 382.474 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 22:43:07 ; Search time 483.789 Seconds  
(without alignments)

1234.209 Million cell updates/sec

Title: US-09-684-215B-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGGGGFAIPI.....QTSKGGTRTGNVILAGPPA 132

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Xgapop 10.0 , Xgapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09684215/runat\_29042004\_061306\_13266/app\_query.fasta.1.1180  
-DB=PublishedApplications NA -QFM=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=ablnu62  
-TRANS=human40.cdi -LIST=45 -LOCAL -OUTFMT=ptc -THR SCORE=100  
-THR MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09684215 @CGN 1.1 370 @runat\_29042004\_061306\_13266  
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US02\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

# RESULT 1

US-09-287-849-27

; Sequence 27, Application US/09287849

; Patent No. US20020009459A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Alderson, Mark

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

; TITLE OF INVENTION: and Their Uses

; FILE REFERENCE: 014058-009020US

; CURRENT APPLICATION NUMBER: US/09/287,849

; CURRENT FILING DATE: 1999-04-07

; PRIOR APPLICATION NUMBER: US 08/818,112

; PRIOR FILING DATE: 1997-03-13

; PRIOR APPLICATION NUMBER: US 08/942,578

; PRIOR FILING DATE: 1997-10-01

; PRIOR APPLICATION NUMBER: US 09/025,197

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 09/056,556

## ALIGNMENTS

1	675	100.0	702	9	US-09-287-849-27	Sequence 27, Appl
2	675	100.0	702	15	US-10-359-460-27	Sequence 27, Appl
3	675	100.0	1002	15	US-10-098-732A-3	Sequence 3, Appl
4	675	100.0	1002	15	US-10-098-732A-5	Sequence 5, Appl
5	675	100.0	1068	9	US-09-712-363-15	Sequence 15, Appl
6	675	100.0	2181	16	US-10-369-983-1	Sequence 1, Appl
7	675	100.0	2190	15	US-10-098-732A-17	Sequence 17, Appl
8	675	100.0	2287	9	US-09-287-849-1	Sequence 1, Appl
9	675	100.0	2287	15	US-10-359-460-1	Sequence 15, Appl
10	675	100.0	2287	15	US-10-098-732A-15	Sequence 15, Appl
11	675	100.0	2287	16	US-10-359-459-1	Sequence 1, Appl
12	675	100.0	2451	16	US-10-369-983-8	Sequence 8, Appl
13	675	100.0	2487	16	US-10-369-983-7	Sequence 7, Appl
14	675	100.0	2637	16	US-10-098-732A-64	Sequence 64, Appl
15	675	100.0	2808	16	US-10-369-983-5	Sequence 5, Appl
16	675	100.0	3030	16	US-10-369-983-3	Sequence 3, Appl
17	675	100.0	3060	16	US-10-369-983-11	Sequence 11, Appl
18	675	100.0	3104	16	US-10-369-983-10	Sequence 10, Appl
19	675	100.0	3474	16	US-10-369-983-9	Sequence 9, Appl
20	675	100.0	447	15	US-10-193-002-4	Sequence 4, Appl
21	670	99.3	447	15	US-10-084-843-4	Sequence 4, Appl
22	670	99.3	447	15	US-10-098-732A-9	Sequence 9, Appl
23	670	99.3	1872	15	US-10-193-002-17	Sequence 17, Appl
24	670	99.3	1872	15	US-10-084-843-17	Sequence 17, Appl
25	670	99.3	1872	15	US-10-098-732A-1	Sequence 1, Appl
26	670	99.3	675	9	US-09-759-143-822	Sequence 822, Appl
27	632	93.6	675	9	US-09-780-669-822	Sequence 822, Appl
28	632	93.6	675	9	US-09-822-827-822	Sequence 822, Appl
29	632	93.6	675	9	US-09-895-793-822	Sequence 822, Appl
30	632	93.6	675	9	US-09-895-814-822	Sequence 822, Appl
31	632	93.6	675	14	US-10-012-896-822	Sequence 822, Appl
32	632	93.6	675	15	US-10-144-678A-822	Sequence 822, Appl
33	632	93.6	675	15	US-10-294-025-822	Sequence 822, Appl
34	632	93.6	822	9	US-09-736-457-1862	Sequence 1862, Ap
35	632	93.6	822	9	US-09-902-941-1862	Sequence 1862, Ap
36	632	93.6	822	9	US-09-849-626-1862	Sequence 1862, Ap
37	632	93.6	822	13	US-10-283-017-1862	Sequence 1862, Ap
38	632	93.6	822	13	US-10-017-754-1862	Sequence 1862, Ap
39	632	93.6	822	15	US-10-113-872-1862	Sequence 1862, Ap
40	632	93.6	894	13	US-09-878-722-236	Sequence 236, App
41	632	93.6	894	13	US-09-904-456-236	Sequence 236, App
42	632	93.6	900	9	US-09-735-705-353	Sequence 353, App
43	632	93.6	900	9	US-09-850-716A-353	Sequence 353, App
44	632	93.6	900	9	US-09-897-778-353	Sequence 353, App
45	632	93.6				

```

; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ral2-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.: 8,198-71 Length: 702
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DBs: 9 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-287-849-27 (1-702)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 25 ACGGCGCGGTCCGATAACTTCCAGCTGTCCAGGTTGGGAGGATTCGCATTCCGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 85 GGGCAGGCGATGCGGATCGCGGCCAGATCCGATCGGTTGGGGGTCAACCACCGTTCA 144
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAnglyAlaArg 60
Db 145 ATCGGGCTACCGCTTCTCCGTGGTGGTGTGTCACAAACACGCAACGGCGCACGA 204
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 205 GTCCAAACGGTGTTCGGAGCGCTCCGGCGCAGTCTCGCATCTCCACCGCGACGTG 264
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 265 ATCACCGCGTGCAGCGGCTCCGATCAATCGGCCACCGCATGGCGGACGGCTTAA 324
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlysserGlyGlyThrArg 120
Db 325 GGGCATATCCCGGTGACGTCTATCTGGTGACCTGGCAACCAACGATCGGGCGCACGCT 384
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 385 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 420

RESULT 2
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio

```

QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCGGTGACGTCACTCGGTGACTGGCAACCAAGTCGGGGCGCACGCGT 384  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTGGCGGAGGAGCCCGCCGCC 420

## RESULT 3

US-10-098-732A-3  
; Sequence 3, Application US/10098732A  
; Publication No. US20030175294A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Brannon, Mark  
; APPLICANT: Guderian, Jeffrey  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
; FILE REFERENCE: 014058-012010US  
; CURRENT APPLICATION NUMBER: US/10/098,732A  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1002  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Ra35 mature  
US-10-098-732A-3

## Alignment Scores:

Pred. No.: 1,26e-70 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-098-732A-3 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 598 ACGGCGGCTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGGCATTCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 658 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 717  
QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
Db 718 ATCGGGCTACCGCTTCTCGGTGGGTGTTCGACAAACACGGCAACGGCGCACGA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 778 GTCCACCGGTGGTGGAGCGCTCCGGTGGGTGTTCGACAAACACGGCAACGGCGCACGA 837  
QY 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 838 ATCAGCGCGTTCGAGCGGCTCCGATCAACTCGGCACCGGATGCGGACGCGCTTAAC 897  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 898 GGGCATCATCCGGTGACGTCACTCGGTGACTGGCAACCAAGTCGGGGCGCACGCGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 958 ACAGGGAACGTGACATTGGCGGAGGAGCCCGCCGCC 993

## RESULT 4

US-10-098-732A-5  
; Sequence 5, Application US/10098732A  
; Publication No. US20030175294A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Brannon, Mark  
; APPLICANT: Guderian, Jeffrey  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
; FILE REFERENCE: 014058-012010US  
; CURRENT APPLICATION NUMBER: US/10/098,732A  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1002  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Ra35FLMutSA  
US-10-098-732A-5

## Alignment Scores:

Pred. No.: 1,26e-70 Length: 1002  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-098-732A-5 (1-1002)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 598 ACGGCGGCTCCGATAACTTCCAGTGTCCAGGGTGGGAGGATTCGGCATTCGATC 657  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 658 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGTGGGGGTCAACCCACCGTTTCA 717  
QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
Db 718 ATCGGGCTACCGCTTCTCGGTGGGTGTTCGACAAACACGGCAACGGCGCACGA 777  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 778 GTCCACCGGTGGTGGAGCGCTCCGGTGGGTGTTCGACAAACACGGCGCACGCGT 837  
QY 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 838 ATCAGCGCGTTCGAGCGGCTCCGATCAACTCGGCACCGGATGCGGACGCGCTTAAC 897  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 898 GGGCATCATCCGGTGACGTCACTCGGTGACTGGCAACCAAGTCGGGGCGCACGCGT 957  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 958 ACAGGGAACGTGACATTGGCGGAGGAGCCCGCCGCC 993

## RESULT 5

US-09-712-363-15  
; Sequence 15, Application US/09712363  
; Patent No. US20020164588A1  
; GENERAL INFORMATION:  
; APPLICANT: Eisenberg, David  
; APPLICANT: Rotstein, Sergio H.  
; APPLICANT: Marcotte, Edward M.  
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
; INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS



;; PRIOR FILING DATE: 2001-03-13  
;; NUMBER OF SEQ ID NOS: 80  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 17  
;; LENGTH: 2190  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence:MTB72MutSA  
;; OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)  
US-10-098-732A-17

Alignment Scores:  
Pred. No.: 3,23e-70 Length: 2190  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-098-732A-17 (1-2190)

Qy 1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGCGCGGTCGGAATACTCCAGCTGCCAGGGTGGCGAGGATTCCGCAATCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGATGGCGATCGCGGCGCATCGATCGGTGGGGGTCCACCACCGTTTCAT 141  
Qy 41 IleGlyProThrAlaPheLeuGlyGlyValValAspAsnAenglyAlaArg 60  
Db 142 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTCGACAAACACGCGACGCGACGA 201  
Qy 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 202 GTCCAAACGCGTGGTGGAGCGCTCCGCGCGCAAGTCTCGCATCTCCACGCGGACGTG 261  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 262 ATCACCAGCGTGGTGGAGCGCTCCGATCACTCGCGCACCGCATGGCGGACGCGCTTAAC 321  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleValSerGlyGlyThrArg 120  
Db 322 GGGCATCATCCGGTGGAGCTATCTCGGTGACCTGGCAAACTCGGCGGCGGCGCGGT 381  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 382 ACAGGAAACGTGACATTGGCGAGGGACCCCGGGCC 417

## RESULT 8

US-09-287-849-1  
; Sequence 1, Application US/09287849  
; Patent No. US2002009459A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556

;; PRIOR FILING DATE: 1998-04-07  
;; PRIOR APPLICATION NUMBER: US 09/223,040  
;; PRIOR FILING DATE: 1998-12-30  
;; NUMBER OF SEQ ID NOS: 46  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 1  
;; LENGTH: 2287  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion  
;; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated MtB32-Mtb39  
;; OTHER INFORMATION: fusion)  
;; NAME/KEY: modified\_base  
;; LOCATION: (30)  
;; OTHER INFORMATION: n = g, a, c or t  
;; NAME/KEY: modified\_base  
;; LOCATION: (33)  
;; OTHER INFORMATION: n = g, a, c or t  
;; NAME/KEY: CDS  
;; LOCATION: (42)..(2231)  
;; NAME/KEY: modified\_base  
;; LOCATION: (2270)  
;; OTHER INFORMATION: n = g, a, c or t  
US-09-287-849-1

Alignment Scores:  
Pred. No.: 3,41e-70 Length: 2287  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-287-849-1 (1-2287)

Qy 1 ThrAlaIaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 63 ACGCGCGGTCGGAATACTCCAGCTGCCAGGGTGGCGAGGATTCCGCAATCCGATC 122  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 123 GGGCAGCGATGGCGATCGCGGCGCATCGCATCGGTGGGGGTCCACCACCGTTTCAT 182  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 183 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTCGACAAACACGCGACGCGACGA 242  
Qy 61 ValGlnArgValValGlySerAlaProAlaIaSerLeuGlyIleSerThrGlyAspVal 80  
Db 243 GTCCAAACGCGTGGTGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGACGTG 302  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 303 ATCACCAGCGTGGAGCGCTCCGATCAACTCGGCGCACCGCATGGCGGACGCGCTTAAC 362  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleValSerGlyGlyThrArg 120  
Db 363 GGGCATCATCCCGTGGAGCTATCTCGGTGACCTGGCAAACTCGGCGGCGGCGCGGT 422  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 423 ACAGGAAACGTGACATTGGCGAGGGACCCCGGGCC 458

## RESULT 9

US-10-359-460-1  
; Sequence 1, Application US/10359460  
; Publication No. US2003014791A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark

APPLICANT: Campos-Neto, Antonio  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
TITLE OF INVENTION: and Their Uses  
FILE REFERENCE: 014058-009020US  
CURRENT APPLICATION NUMBER: US/10/359,460  
CURRENT FILING DATE: 2003-02-05  
PRIOR APPLICATION NUMBER: US/09/287,849  
PRIOR FILING DATE: 1999-04-07  
PRIOR APPLICATION NUMBER: US 08/818,112  
PRIOR FILING DATE: 1997-03-13  
PRIOR APPLICATION NUMBER: US 08/942,578  
PRIOR FILING DATE: 1997-10-01  
PRIOR APPLICATION NUMBER: US 09/025,197  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 09/056,556  
PRIOR FILING DATE: 1998-04-07  
PRIOR APPLICATION NUMBER: US 09/223,040  
PRIOR FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2287  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
OTHER INFORMATION: Protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39  
OTHER INFORMATION: fusion)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (30)  
OTHER INFORMATION: n = g, a, c o r t  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (33)  
OTHER INFORMATION: n = g, a, c o r t  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (42)..(2231)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (2270)  
OTHER INFORMATION: n = g, a, c o r t  
US-10-359-460-1  
Alignment Scores:  
Pred. No.: 3.41e-70 Length: 2287  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0  
US-09-684-215B-4 (1-132) x US-10-359-460-1 (1-2287)  
QY 1 ThrAlaAlaSerAspAsnGlnGlyGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
Db 63 ACGCCGCGTCCGATTAATCCAGCTGCCAGGTGGGAGGATTCGCCATTCGATC 122  
QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 123 GGGCAGCGCATGCGATCGCGGCGCAGATCCGATCGGTGGGGGTCACCACCGTTCAT 182  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlnGlyValAlaArg 60  
Db 183 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTGTGACAAACACGCGACGCGACGA 242  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 243 GTCCACGCGTGTGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 302  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 303 ATCACCGCGTGCAGCGGCTCCGATCACTCGGCCACCGCATGCGCGACCGCTTAAC 362  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 363 GGGCATCATCCCGGTGACGTCTCGGTGACCTCGCAACCAAGTCGGCGGCGCGGT 422  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 423 ACAGGAGACGTGACATTCGCCGAGGAGACCCCGGCC 458  
RESULT 10  
US-10-098-732A-15  
Sequence 15, Application US/10098732A  
Publication No. US20030175294A1  
GENERAL INFORMATION:  
APPLICANT: Skeiky, Yasir  
APPLICANT: Brannon, Mark  
APPLICANT: Guderian, Jeffrey  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
TITLE OF INVENTION: Leishmania Antigen  
FILE REFERENCE: 014058-012010US  
CURRENT APPLICATION NUMBER: US/10/098,732A  
CURRENT FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: US 60/275,837  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 15  
LENGTH: 2287  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
OTHER INFORMATION: protein Mtb72F (Ra12-TbH9-Ra35 or Mtb32-Mtb39  
OTHER INFORMATION: fusion)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (42)..(2231)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1)..(2287)  
OTHER INFORMATION: n = g, a, c o r t  
US-10-098-732A-15  
Alignment Scores:  
Pred. No.: 3.41e-70 Length: 2287  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0  
US-09-684-215B-4 (1-132) x US-10-098-732A-15 (1-2287)  
QY 1 ThrAlaAlaSerAspAsnGlnGlyGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
Db 63 ACGCCGCGTCCGATTAATCCAGCTGCCAGGTGGGAGGATTCGCCATTCGATC 122  
QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 123 GGGCAGCGCATGCGATCGCGGCGCAGATCCGATCGGTGGGGGTCACCACCGTTCAT 182  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlnGlyValAlaArg 60  
Db 183 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTGTGACAAACACGCGACGCGACGA 242  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 243 GTCCACGCGTGTGGGAGCGTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 302

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 303 ATCCGCGGTCGACGGCGCTCCGATCACTCGGCCACCGGATCGCGGACGGCTTAAC 362  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleArgSerGlyThrArg 120  
 Db 363 GGGCATCATCCGGTGACGTCTATCGGTGACCTGGCAACCAAGTCGGCGGCACGCGT 422  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 423 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 458

RESULT 11  
 US-10-359-459-1  
 ; Sequence 1, Application US/10359459  
 ; Publication No. US20040013677A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; FILE REFERENCE: 014058-009010US  
 ; CURRENT APPLICATION NUMBER: US/10/359,459  
 ; PRIOR FILING DATE: 2003-02-05  
 ; PRIOR APPLICATION NUMBER: US/09/223,040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 2287  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
 ; OTHER INFORMATION: protein Ra12-TbH9-Ra35  
 ; NAME/KEY: modified\_base  
 ; LOCATION: (30)  
 ; OTHER INFORMATION: n = g, a, c or t  
 ; NAME/KEY: modified\_base  
 ; LOCATION: (33)  
 ; OTHER INFORMATION: n = g, a, c or t  
 ; NAME/KEY: CDS  
 ; LOCATION: (42)..(2231)  
 ; FEATURE:  
 ; NAME/KEY: modified\_base  
 ; LOCATION: (2270)  
 ; OTHER INFORMATION: n = g, a, c or t  
 ; US-10-359-459-1

Alignment Scores:  
 Pred. No.: 3,41e-70 Length: 2287  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-359-459-1 (1-2287)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 63 ACGGCGCGTCCGATCACTTCAGCTGTCCAGGCTGGCGAGGATTCGCCATTCCGATC 122  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 123 GGGCAGCGGATCGCATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCCGTTTCA 182  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60

Db 183 ATCCGCGCTACCGCTTCTCGCTTGGGTGTGTGCAACAACGCAACCGCGCACGA 242  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 243 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAATCTCGCATTTCCACCGCGCACGTG 302  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 303 ATCCGCGGTCGACGGCGCTCCGATCACTCGGCCACCGGATCGCGGACGGCTTAAC 362  
 QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleArgSerGlyThrArg 120  
 Db 363 GGGCATCATCCGGTGACGTCTATCGGTGACCTGGCAACCAAGTCGGCGGCACGCGT 422  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 423 ACAGGGAACGTGACATTGGCCGAGGACCCCGGCC 458

RESULT 12  
 US-10-369-983-8  
 ; Sequence 8, Application US/10369983  
 ; Publication No. US20030235593A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Guderian, Jeff  
 ; APPLICANT: Reed, Steven  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 ; FILE REFERENCE: 014058-009081US  
 ; CURRENT APPLICATION NUMBER: US/10/369,983  
 ; CURRENT FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/357,351  
 ; PRIOR FILING DATE: 2002-02-15  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 2451  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; OTHER INFORMATION: MTB81F (MTB72F-DPV)  
 ; US-10-369-983-8

Alignment Scores:  
 Pred. No.: 3,71e-70 Length: 2451  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-8 (1-2451)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 25 ACGGCGCGTCCGATCACTTCAGCTGTCCAGGCTGGCGAGGATTCGCCATTCCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 85 GGGCAGCGGATCGCATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCCGTTTCA 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
 Db 145 ATCCGGCGCTACCGCTTCTCGGTGGGTGTGTCACAACAACGCAACCGCGCACGA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 205 GTCCAAACGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGCATTTCCACCGCGCACGTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100



Db 265 ATCACCGCGTGCAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGGCGTTTAAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 420

## RESULT 13

US-10-369-983-7  
; Sequence 7, Application US/10369983  
; Publication No. US20030235593A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderian, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; PRIOR FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 2487  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
; OTHER INFORMATION: MTH83F (MTH72F-MTI)  
US-10-369-983-7

Alignment Scores:  
Pred. No.: 3,77e-70 Length: 2487  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-7 (1-2487)

Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 25 ACAGCGCGTCCGATCAACTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 84  
Qy 21 GlyAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGCATGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCCGTTTCAAT 144  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 145 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTCGACCAACACCGGACCGGACGCA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCCACGCGTGTGGAGCGCTCCGGGGCAGATCCGATCGGTGGGGGTCAACCCCGTTTCAAT 264  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACCGCGTGCAGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTAAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 420

## RESULT 14

US-10-369-983-6  
; Sequence 6, Application US/10369983  
; Publication No. US20030235593A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Guderian, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT APPLICATION NUMBER: US/10/369,983  
; PRIOR FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US 60/357,351  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 2637  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
; OTHER INFORMATION: MTH89F (MTH72F-Erd14)  
US-10-369-983-6

Alignment Scores:  
Pred. No.: 4,05e-70 Length: 2637  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-6 (1-2637)

Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 25 ACAGCGCGTCCGATCAACTTCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 84  
Qy 21 GlyAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGCATGCGATCGCGGCCAGATCCGATCGGTGGGGGTCAACCCCGTTTCAAT 144  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 145 ATCGGGCTACCGCTTCCCTCGGCTTGGGTGTTCGACCAACACCGGACCGGACGCA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 205 GTCCACGCGTGTGGAGCGCTCCGGGGCAGATCCGATCGGTGGGGGTCAACCCCGTTTCAAT 264  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 265 ATCACCGCGTGCAGCGCGTCCGATCAACTCGGCCACCGCGATGGCGGACGGCTTAAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCGGGGCGGACGGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 420

## RESULT 15

US-10-098-732A-64  
; Sequence 64, Application US/10098732A  
; Publication No. US20030175294A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Brannon, Mark  
; APPLICANT: Guderian, Jeffrey

```

; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; FILE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 2808
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72F-MAPS
; OTHER INFORMATION: (r95f) fusion construct, TB MTB72F (Ra12-TbH9-Ra35)
; OTHER INFORMATION: linked to Leishmania thiol-specific antioxidant
; OTHER INFORMATION: (TSA or MAPS)
US-10-098-732A-64

Alignment Scores:
Pred. No.: 4,37e-70 Length: 2808
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-098-732A-64 (1-2808)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 25 ACGGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGGAGGATTCGCCATTCGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 85 GGGCAGGCGATCGCATCCGCGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTTCAT 144
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60
Db 145 ATCGGGCGCTACCGCTTCCTCGCTTGGGTGGTGTTCGACCAACAGCGGCAACGGCGCAGCA 204
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyVileSerThrGlyAspVal 80
Db 205 GTTCAACCGTGTGTCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCAGGTG 264
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 265 ATCACCAGCGGTGCGAGCGCTCCGATCACTCGGCCACCGCATGGCGGACCGCTTAAC 324
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 325 GGGCATCATCCCGGTGACGTCACTCTCGGTGACCTGGCAACCAAGTCCGGCGGACCGGT 384

RESULT 16
US-10-369-983-5
; Sequence 5, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3030
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion

```

```

; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2808
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-5

Alignment Scores:
Pred. No.: 4,37e-70 Length: 2808
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-5 (1-2808)
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db 25 ACGGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGGAGGATTCGCCATTCGATC 84
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 85 GGGCAGGCGATCGCATCCGCGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTTCAT 144
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60
Db 145 ATCGGGCGCTACCGCTTCCTCGCTTGGGTGGTGTTCGACCAACAGCGGCAACGGCGCAGCA 204
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyVileSerThrGlyAspVal 80
Db 205 GTTCAACCGTGTGTCGGAGCGCTCCGCGGCGCAAGTCTCGGCATCTCCACCGCGCAGGTG 264
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 265 ATCACCAGCGGTGCGAGCGCTCCGATCACTCGGCCACCGCATGGCGGACCGCTTAAC 324
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 325 GGGCATCATCCCGGTGACGTCACTCTCGGTGACCTGGCAACCAAGTCCGGCGGACCGGT 384

RESULT 17
US-10-369-983-3
; Sequence 3, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3030
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion

```

; OTHER INFORMATION: protein

US-10-369-983-3

Alignment Scores:

Pred. No.: 4,79e-70 Length: 3030  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-3 (1-3030)

QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 595 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGGTGGGACAGGATTCGCCATTCCGATC 654  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 DB 655 GGGCAGGCGATCGCATCGCGGCGAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAAT 714  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyValAlaArg 60  
 DB 715 ATCGGCGCTACCGCTTCTCGCTTGGTGTGTGCGACAACACGCGCAACGGGCGACGA 774  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 775 GTCCAAACGCGTGGTGGGAGCGCTCGGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 834  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 835 ATCAGCGCGTGCAGCGGCTCCGATCACTGGCCACCGCATGGCGGACGCGCTTAAC 894  
 QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 DB 895 GGGCATCATCCCGGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGGT 954  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 955 ACAGGGAACGTGACATGGCCGAGGACCCCGGCC 990

RESULT 18

US-10-369-983-11

; Sequence 11, Application US/10369983

; Publication No. US20030235593A1

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir

; APPLICANT: Gudarian, Jeff

; APPLICANT: Reed, Steven

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis

; FILE REFERENCE: 014058-009081US

; CURRENT APPLICATION NUMBER: US/10/369,983

; CURRENT FILING DATE: 2003-02-18

; PRIOR APPLICATION NUMBER: US 60/357,351

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 11

; LENGTH: 3060

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: fusion protein

; OTHER INFORMATION: MTB103F (MTB72F-85b)

US-10-369-983-11

Alignment Scores:

Pred. No.: 4,84e-70 Length: 3060  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

DB:

16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-11 (1-3060)

QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 25 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGGTGGGACAGGATTCGCCATTCCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 DB 85 GGGCAGGCGATGGCATCGCGGCGAGATCCGATCGGGTGGGGGTCAACCACCGTTTCAAT 144  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
 DB 145 ATCGGCGCTACCGCTTCTCGCTTGGTGTGTGCGACAACACGCGCAACGGGCGACGA 204  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 205 GTCCAAACGCGTGGTGGGAGCGCTCGGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 264  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 265 ATCAGCGCGTGCAGCGGCTCCGATCACTGGCCACCGCATGGCGGACGCGCTTAAC 324  
 QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 DB 325 GGGCATCATCCCGGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGGT 384  
 QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 DB 385 ACAGGGAACGTGACATGGCCGAGGACCCCGGCC 420

RESULT 19

US-10-369-983-10

; Sequence 10, Application US/10369983

; Publication No. US20030235593A1

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir

; APPLICANT: Gudarian, Jeff

; APPLICANT: Reed, Steven

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis

; FILE REFERENCE: 014058-009081US

; CURRENT APPLICATION NUMBER: US/10/369,983

; CURRENT FILING DATE: 2003-02-18

; PRIOR APPLICATION NUMBER: US 60/357,351

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 10

; LENGTH: 3104

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: fusion protein

; OTHER INFORMATION: MTB102tm2F (MTB102FTM, MTB72F-hTCC#1)

US-10-369-983-10

Alignment Scores:

Pred. No.: 4,93e-70 Length: 3104  
 Score: 675.00 Matches: 132  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-369-983-10 (1-3104)

QY 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 25 ACGGCGCGTCCGATAAATCCAGCTGTCCAGGGTGGGACAGGATTCGCCATTCCGATC 84  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

Db 85 GGGCAGCGGATGGCGATCGCGGCGCAGATCCGATCGGTTGGGGGTCCACCCACCGTTTCAT 144  
Qy 41 IleglyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValArg 60  
Db 145 ATCGGGCTACCGCTTCTCGCTTGGTGTTCGACACACACGCGACGCGCACGA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyLeuGlySerThrGlyAspVal 80  
Db 205 GTCCAAACGCGTGTTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 264  
Qy 81 IleglyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValArg 100  
Db 265 ATCACCAGCGGTGCGACGCGCTCCGATCACTCGGCACCGCGATGGCGACGCGCTTAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTGGGGGGCACGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTCGCGGAGGACCCCGCGCC 420  
RESULT 20  
US-10-369-983-9  
; Sequence 9, Application US/10369983  
; Publication No. US20030235593A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Gudarian, Jeff  
; APPLICANT: Reed, Steven  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
; FILE REFERENCE: 014058-009081US  
; CURRENT FILING DATE: 2003-02-18  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 3474  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
; OTHER INFORMATION: MTB114F (MTB72F-mTCC#2)  
US-10-369-983-9  
Alignment Scores:  
Pred. No.: 5.65e-70 Length: 3474  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 16 Gaps: 0  
US-09-684-215B-4 (1-132) x US-10-369-983-9 (1-3474)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 25 ACAGCGCGCTCGATTAACCTCCAGCTGTCCAGGTGGCGAGGATTCGCCATTCGATC 84  
Qy 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 85 GGGCAGCGCATCGCATCGCGGCCAGATCCGATCGGTTGGGGGTCCACCCACCGCTTCAT 144  
Qy 41 IleglyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValArg 60  
Db 145 ATCGGGCTACCGCTTCTCGCTTGGTGTTCGACACACACGCGACGCGCACGA 204  
Qy 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyLeuGlySerThrGlyAspVal 80

Db 205 GTCCAAACGCGTGTTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264  
Qy 81 IleglyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyValArg 100  
Db 265 ATCACCAGCGGTGCGACGCGCTCCGATCACTCGGCACCGCGATGGCGACGCGCTTAAC 324  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 325 GGGCATCATCCCGGTGACGTCTCTCGGTGACCTGGCAACCAAGTGGGGGGCACGCGT 384  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 385 ACAGGGAACGTGACATTCGCGGAGGACCCCGCGCC 420  
RESULT 21  
US-10-193-002-4  
; Sequence 4, Application US/10193002  
; Publication No. US20030135026A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/193,002  
; FILING DATE: 10-Jul-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,596  
; FILING DATE: 05-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-193-002-4  
Alignment Scores:  
Pred. No.: 1.86e-70 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 15 Gaps: 0

```
US-09-684-215B-4 (1-132) x US-10-193-002-4 (1-447)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCCGGTCCGATAAATTCAGCTGTCACAGGTTGGCAGGATTCGCAATCCGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCGGTTGGGGGTCCACCCACCGTTTCAT 130
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60
Db 131 ATCGGGCTACCGCTTCTCGGTTGGTGTTCACACACACAGCGGCGGCGACGA 190
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGCGTGGTGGGAGGCTCCGGCGGAGATCCGATCCGATCCGCGGCGAGGTG 250
QY 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerIleAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCCCGGGTCCGCGGCTCCGATCAATCGGCGACCGGATGGCGGCGCGCTTAAAC 310
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 311 GGGCATCATCCCGGTGACGTATCTCGGTGAATGCGCAACCAAGTCGGGCGGCGACGCT 370
QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 371 ACAGGAACGTGACATTCGCGGAGGACCCCGGCC 406

RESULT 22
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; Skeiky, Yasir A.W.
; Dillon, Davin C.
; Campos-Neto, Antonio
; Houghton, Raymond
; Vedvick, Thomas S.
; Twardzik, Daniel R.
; Lodes, Michael J.
; Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
```

```
TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4
Alignment Scores:
Pred. No.: 1-86e-70 Length: 447
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 15 Gaps: 0
US-09-684-215B-4 (1-132) x US-10-084-843-4 (1-447)
QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 11 ACGGCCGGTCCGATAAATTCAGCTGTCACAGGTTGGCAGGATTCGCAATCCGATC 70
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 71 GGGCAGGCGATGGCGATCGCGGCGAGATCCGATCCGATCCGCGGCGAGGTG 130
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60
Db 131 ATCGGGCTACCGCTTCTCGGTTGGTGTTCACACACACAGCGGCGGCGACGA 190
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 191 GTCCAAACGCGTGGTGGGAGGCTCCGGCGGAGATCCGATCCGATCCGCGGCGAGGTG 250
QY 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerIleAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 251 ATCCCGGGTCCGCGGCTCCGATCAATCGGCGACCGGATGGCGGCGCGCTTAAAC 310
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 311 GGGCATCATCCCGGTGACGTATCTCGGTGAATGCGCAACCAAGTCGGGCGGCGACGCT 370
QY 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 371 ACAGGAACGTGACATTCGCGGAGGACCCCGGCC 406

RESULT 23
US-10-098-732A-9
; Sequence 9, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderman, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTBRa12 (MTB32A)
; OTHER INFORMATION: C-terminus)
```

US-10-098-732A-9

Alignment Scores:

Pred. No.:	1.86e-70	Length:	447
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-4 (1-132) x US-10-098-732A-9 (1-447)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProle 20

Db 11 ACGGCGGCGTCCGATTAATCCAGCTTCCAGGTCGGCAGGATTCGCCATTCGGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

Db 71 GGGCAGCGGATCGGATCGCGGCCAGATCCGATCGGTCGGGGGGTCACCCACCGTTTCAT 130

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60

Db 131 ATCGGCGCTACCGCTTCTCGCTTGGGTGTGTGTCACACACACGCGACGCGGACGA 190

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerIleuGlyIleSerThrGlyAspVal 80

Db 191 GTCCAAACGCGTGTCCGCGAGCGCTCCGCGGCGAAGTCTCGGCATCTCCACCGCGACGTG 250

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 251 ATACCGCGGTTCAGCGGCTTCGATCACTCGGCCACCGCGATGCGGCGCTTAAC 310

QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120

Db 311 GGGCATCATCCCGTACGTCATCTCGGTGAACCTGGCAACCAAGTCGGCGCGCACGCGT 370

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 371 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 406

RESULT 24

US-10-193-002-17

Sequence 17, Application US/10193002

Publication No. US20030135026A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

Skeiky, Yasir A.W.

Dillon, Davin C.

Campos-Neto, Antonio

Houghton, Raymond

Vedvick, Thomas S.

Twardzik, Daniel R.

Lodes, Michael J.

Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 350

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/193,002

FILING DATE: 10-Jul-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/072,596

FILING DATE: 05-MAY-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 692-6031

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 1872 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 17:

US-10-193-002-17

Alignment Scores:

Pred. No.:	1.05e-69	Length:	1872
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-4 (1-132) x US-10-193-002-17 (1-1872)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProle 20

Db 758 ACGGCGGCGTCCGATTAATCCAGCTTCCAGGTCGGCAGGATTCGCCATTCGGATC 817

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

Db 818 GGGCAGCGGATCGGATCGCGGCCAAATCCGATCGGTCGGGGGGTCACCCACCGTTTCAT 877

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60

Db 878 ATCGGCGCTACCGCTTCTCGGTGTGTGTCACACACACGCGACGCGGACGA 937

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 938 GTCCAAACGCGTGTCCGGAAGCGTCCGCGCGCAAGTCTCGGCATCTCCACCGCGACGTG 997

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 998 ATACCGCGGTTCAGCGGCTTCGATCACTCGGCCACCGCGATGCGGCGCTTAAC 1057

QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120

Db 1058 GGGCATCATCCCGTACGTCATCTCGGTGAACCTGGCAACCAAGTCGGCGCGCACGCGT 1117

QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 1118 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 1153

RESULT 25

US-10-084-843-17

Sequence 17, Application US/10084843

Publication No. US20030143243A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

Skeiky, Yasir A.W.

Dillon, Davin C.

Campos-Neto, Antonio

Houghton, Raymond

Vedvick, Thomas S.

Twardzik, Daniel R.

Lodes, Michael J.

Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/084,843  
FILING DATE: 25-Feb-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-10-084-843-17

Alignment Scores:  
Pred. No.: 1,05e-69 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservatives: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-084-843-17 (1-1872)

Qy	1	ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	758	ACGCGCGCTCCGATACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC	817
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
Db	818	GGGCAGCGATGGCGATCGGGGCAAAATCCGATCGGGTGGGGGTCACCACCGTTTCAT	877
Qy	41	IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValAsnGlyAlaArg	60
Db	878	ATCGGGCTACCGCTTCTCGCTTGGTGTGTCTCGACACACGCGACGCGACGA	937
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	878	ATCGGGCTACCGCTTCTCGCTTGGTGTGTCTCGACACACGCGACGCGACGA	937
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	938	GTCCACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGACGTG	997
Qy	101	GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
Db	998	ATCACCGGGTCCGCGCTCCGATCAATCCGCCACCGCGATGGCGGACGCGCTTAAC	1057
Qy	101	GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
Db	1058	GGGCATCATCCGGTGGAGTCTTCGGTGAACATGGCAACCAAGTCGGGCGGACGCGT	1117
Qy	121	ThrGlyAsnValThrLeuAlaGlyProProAla	132
Db	1118	ACAGGGAACGTGACATTGGCGAGGAGGCCCGCGCC 1153	

## RESULT 27

US-09-759-143-822  
; Sequence 822, Application US/09759143  
; Patent No. US2002022248A1

## RESULT 26

US-10-098-732A-1  
; Sequence 1, Application US/10098732A  
; Publication No. US20030175294A1  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Brannon, Mark  
; APPLICANT: Gudarian, Jeffrey  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a  
; FILE OF INVENTION: Leishmania Antigen  
; FILE REFERENCE: 014058-012010US  
; CURRENT APPLICATION NUMBER: US/10/098,732A  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: US 60/275,837  
; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1872  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: MTB32A (Ra35FL)  
; FEATURE:  
; NAME/KEY: modified base  
; LOCATION: (1)..(1872)  
; OTHER INFORMATION: n = g, a, c or t  
US-10-098-732A-1

## Alignment Scores:

Pred. No.: 1,05e-69 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservatives: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 15 Gaps: 0

US-09-684-215B-4 (1-132) x US-10-098-732A-1 (1-1872)

Qy	1	ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	758	ACGCGCGCTCCGATACTTCCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC	817
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
Db	818	GGGCAGCGATGGCGATCGGGGCAAAATCCGATCGGGTGGGGGTCACCACCGTTTCAT	877
Qy	41	IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValAsnGlyAlaArg	60
Db	878	ATCGGGCTACCGCTTCTCGCTTGGTGTGTCTCGACACACGCGACGCGACGA	937
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	938	GTCCACCGGTGGTGGAGCGCTCCGGCGCAAGTCTCGCATCTCCACCGCGACGTG	997
Qy	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	998	ATCACCGGGTCCGCGCTCCGATCAATCCGCCACCGCGATGGCGGACGCGCTTAAC	1057
Qy	101	GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
Db	1058	GGGCATCATCCGGTGGAGTCTTCGGTGAACATGGCAACCAAGTCGGGCGGACGCGT	1117
Qy	121	ThrGlyAsnValThrLeuAlaGlyProProAla	132
Db	1118	ACAGGGAACGTGACATTGGCGAGGAGGCCCGCGCC 1153	

GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darriack  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C23  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; CURRENT FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-759-143-822

Alignment Scores:  
Pred. No.: 9,89e-66 Length: 675  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 9 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-759-143-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGCGCGGTCGATACCTCCAGCTGCCAGGTCGGCAGGATTCGCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----CTTCCACCGCTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyValaArg 60  
Db 130 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCGACACACGCGGCGCAGCA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCAAACGCGTGGTCGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGCAGTG 249  
Qy 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGGTCGACGCGCTCCGATCACTCGGCGCACCGGATGGCGGACGGCTTAAC 309  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 310 GGGCATCATCCCGGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGGCACGCGT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC 405

RESULT 28

US-09-780-669-822  
; Sequence 822, Application US/09780669  
; Patent No. US2002005197A1

GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darriack  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT APPLICATION NUMBER: US/09/780,669  
; CURRENT FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 822  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-780-669-822

Alignment Scores:  
Pred. No.: 9,89e-66 Length: 675  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 9 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-780-669-822 (1-675)  
Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACGCGCGGTCGATACCTCCAGCTGCCAGGTCGGCAGGATTCGCATTCGGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----CTTCCACCGCTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyValaArg 60  
Db 130 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCGACACACGCGGACGGCAGCA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCAAACGCGTGGTCGGGAGCGCTCCGCGGCAAGTCTCGCATCTCCACCGCGCAGTG 249  
Qy 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGGTCGACGCGCTCCGATCACTCGGCGCACCGGATGGCGGACGGCTTAAC 309  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 310 GGGCATCATCCCGGTGACGTCTCTCGTGACCTGGCAACCAAGTCCGGCGGCACGCGT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTGGCCGAGGGACCCCGGCC 405

RESULT 29



## US-09-822-827-822

; Sequence 822, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C1

; CURRENT APPLICATION NUMBER: US/09/822.827

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-822-827-822

## Alignment Scores:

Pred. No.: 9,886-66 Length: 675  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservatives: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 9 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-822-827-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGGTCCGATTAATCCAGCTGCCAGGGTGGCGAGGATTCGCATTCGCATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 82 GGGCAGCGGATGCGGATCCGCGGCCAGATCAAG-----CTTCCACCGTTTCAT 129
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60
DB 130 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTGCACAAACACGGCAACGGCGGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 190 GTCCACACGGTGTGGGAGCGCTCCGGCGCAAGTCTCGGCACTCCACCGCGGACGTG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 250 ATCAGCGCGGTCCAGCGGCTCCGATCAACTCGGCCACCGGATGGCGGACGGCGTTAAC 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACCGGT 369
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 370 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 405
```

## RESULT 30

US-09-895-793-822

; Sequence 822, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Devin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yufu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Repler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C2

; CURRENT APPLICATION NUMBER: US/09/895,793

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-793-822

## Alignment Scores:

Pred. No.: 9,886-66 Length: 675  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservatives: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 9 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-895-793-822 (1-675)

```
QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
DB 22 ACGGCGCGGTCCGATTAATCCAGCTGCCAGGGTGGCGAGGATTCGCATTCGCATC 81
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
DB 82 GGGCAGCGGATGCGGATCCGCGGCCAGATCAAG-----CTTCCACCGTTTCAT 129
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAlaArg 60
DB 130 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTGTGCACAAACACGGCAACGGCGGACGA 189
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
DB 190 GTCCACACGGTGTGGGAGCGCTCCGGCGCAAGTCTCGGCACTCCACCGCGGACGTG 249
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
DB 250 ATCAGCGCGGTCCAGCGGCTCCGATCAACTCGGCCACCGGATGGCGGACGGCGTTAAC 309
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
DB 310 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACCGGT 369
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
DB 370 ACAGGGAACGTGACATTGGCCGAGGAGACCCCGGCC 405
```

Search completed: April 30, 2004, 05:10:21

Job time : 493.039 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 22:28:57 ; Search time 67.5789 Seconds  
(without alignments)  
1083.969 Million cell updates/sec

Title: US-09-684-215B-4

Perfect score: 675

Sequence: 1 TAASDNFQLSQGGGFAIP1.....QTKSGGTRTGNVTLAGPPA 132

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/uspto.spool/US09684215/runat 29042004 061306 13249/app\_query.fasta\_1.1180  
-DB=Issued Patents NA -QFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=30  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215 @CGN 1 1 115 @runat 29042004 061306 13249 -NCFU=6 -ICFU=3  
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

- 1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	675	100.0	702	4	US-09-287-849-27 Sequence 27, Appl
2	675	100.0	2287	4	US-09-223-040-1 Sequence 1, Appl
3	675	100.0	2287	4	US-09-287-849-1 Sequence 1, Appl
4	675	100.0	4403765	3	US-09-103-840A-2 Sequence 2, Appl
5	675	100.0	4411529	3	US-09-103-840A-1 Sequence 1, Appl
6	670	99.3	447	3	US-08-818-112-4 Sequence 4, Appl
7	670	99.3	447	4	US-08-818-111-4 Sequence 4, Appl
8	670	99.3	447	4	US-09-056-556-4 Sequence 4, Appl
9	670	99.3	447	4	US-09-072-967-4 Sequence 4, Appl
10	670	99.3	447	4	US-09-072-967-4 Sequence 4, Appl
11	670	99.3	1872	3	US-08-818-112-17 Sequence 17, Appl
12	670	99.3	1872	4	US-08-818-111-17 Sequence 17, Appl

13	670	99.3	1872	4	US-09-056-556-17 Sequence 17, Appl
14	670	99.3	1872	4	US-09-072-967-17 Sequence 17, Appl
15	670	99.3	1872	4	US-09-072-967-17 Sequence 17, Appl
16	632	93.6	675	4	US-09-636-215-822 Sequence 822, App
17	632	93.6	675	4	US-09-636-215-822 Sequence 822, App
18	632	93.6	822	4	US-09-736-457-1862 Sequence 1862, Ap
19	632	93.6	900	4	US-09-643-597-353 Sequence 353, App
20	632	93.6	900	4	US-09-606-421B-353 Sequence 353, App
21	632	93.6	915	4	US-09-636-215-834 Sequence 834, App
22	632	93.6	915	4	US-09-685-166A-834 Sequence 834, App
23	632	93.6	945	4	US-09-736-457-1861 Sequence 1861, Ap
24	632	93.6	1012	4	US-09-643-597-351 Sequence 351, App
25	632	93.6	1012	4	US-09-606-421B-351 Sequence 351, App
26	632	93.6	1203	4	US-09-636-215-851 Sequence 851, App
27	632	93.6	1203	4	US-09-685-166A-851 Sequence 851, App
28	632	93.6	1464	4	US-09-620-412C-348 Sequence 348, App
29	632	93.6	1464	4	US-09-598-419-348 Sequence 348, App
30	632	93.6	1557	4	US-09-620-412C-332 Sequence 332, App
31	632	93.6	1557	4	US-09-598-419-332 Sequence 332, App
32	632	93.6	1578	4	US-09-556-877-188 Sequence 188, App
33	632	93.6	1578	4	US-09-620-412C-188 Sequence 188, App
34	632	93.6	1578	4	US-09-598-419-188 Sequence 188, App
35	632	93.6	1752	4	US-09-620-412C-352 Sequence 352, App
36	632	93.6	1752	4	US-09-598-419-352 Sequence 352, App
37	632	93.6	1758	4	US-09-620-412C-336 Sequence 336, App
38	632	93.6	1758	4	US-09-598-419-336 Sequence 336, App
39	632	93.6	1860	4	US-09-620-412C-308 Sequence 308, App
40	632	93.6	1860	4	US-09-598-419-308 Sequence 308, App
41	632	93.6	1896	4	US-09-620-412C-324 Sequence 324, App
42	632	93.6	1896	4	US-09-598-419-324 Sequence 324, App
43	632	93.6	1941	4	US-09-620-412C-316 Sequence 316, App
44	632	93.6	1941	4	US-09-598-419-316 Sequence 316, App
45	632	93.6	1965	4	US-09-620-412C-340 Sequence 340, App

ALIGNMENTS

RESULT 1

US-09-287-849-27  
; Sequence 27, Application US/09287849  
; Patent No. 6627198  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Alderson, Mark  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
; TITLE OF INVENTION: and Their Uses  
; FILE REFERENCE: 014058-009020US  
; CURRENT APPLICATION NUMBER: US/09/287,849  
; CURRENT FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: US 08/818,112  
; PRIOR FILING DATE: 1997-03-13  
; PRIOR APPLICATION NUMBER: US 08/942,578  
; PRIOR FILING DATE: 1997-10-01  
; PRIOR APPLICATION NUMBER: US 09/025,197  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 702  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion  
; OTHER INFORMATION: protein Rai2-DPPD (designated Mtb24), reading  
; OTHER INFORMATION: frame 1

```
; NAME/KEY: CDS
; LOCATION: (1)..(693)
; OTHER INFORMATION: bi-fusion protein Ral2-DPPD (designated Mtb24),
; OTHER INFORMATION: reading frame 1
; NAME/KEY: CDS
; LOCATION: (2)..(700)
; OTHER INFORMATION: reading frame 2
; NAME/KEY: CDS
; LOCATION: (3)..(701)
; OTHER INFORMATION: reading frame 3
US-09-287-849-27
```

```
Alignment Scores:
Pred. No.: 2,076-65 Length: 702
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
```

US-09-684-215B-4 (1-132) x US-09-287-849-27 (1-702)

```
Qy 1 ThrAlaAlaSerAspAsnPhedInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCCGATC 84
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 85 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 144
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 145 ATCGGGCCATCGCCCTTCTCGGCTTGGGTGTGTGCGAACAAACGGCGACGGCGACGA 204
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 205 GTCCAAACGGGTGTCGGAGGCGCTCCGGGGCACTCTCGCATCTCCACCGGCGAGGTG 264
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 265 ATCAGCGGGTCCAGCGGCTCCGATCACTCGGCCACCGGATGGCGAGCGCGCTTAAC 324
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 325 GGGCATCATCCCGGTGAGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACCGGT 384
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 385 ACAGGGAACGTGACATTGGCGAGGACCCCGGGCC 420
```

#### RESULT 2

```
US-09-223-040-1
; Sequence 1, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/09/223,040
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
; OTHER INFORMATION: Protein Ral2-TbH9-Ra35
```

```
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-09-223-040-1
```

```
Alignment Scores:
Pred. No.: 9,786-65 Length: 2287
Score: 675.00 Matches: 132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
```

US-09-684-215B-4 (1-132) x US-09-223-040-1 (1-2287)

```
Qy 1 ThrAlaAlaSerAspAsnPhedInLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGCGAGGATTCCGCATTCCGATC 122
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 123 GGGCAGGCGATGGCGATCGGGGCCAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCA 182
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 183 ATCGGGCCATCGCCCTTCTCGGCTTGGGTGTGTGCGAACAAACGGCGACGGCGACGA 242
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCAAACGGGTGTCGGAGGCGCTCCGGGGCACTCTCGCATCTCCACCGGCGAGGTG 302
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCAGCGGCTCGACGCGCTCCGATCACTCGGCCACCGCGATGGCGAGCGCGCTTAAC 362
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 363 GGGCATCATCCCGGTGAGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACCGGT 422
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 423 ACAGGGAACGTGACATTGGCGAGGACCCCGGGCC 458
```

#### RESULT 3

```
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
```

; PRIOR APPLICATION NUMBER: US 09/056,556  
; PRIOR FILING DATE: 1998-04-07  
; PRIOR APPLICATION NUMBER: US 09/223,040  
; PRIOR FILING DATE: 1998-12-30  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2287  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: tri-fusion  
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39  
; OTHER INFORMATION: fusion)  
; NAME/KEY: modified\_base  
; LOCATION: (30)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: modified\_base  
; LOCATION: (33)  
; OTHER INFORMATION: n = g, a, c or t  
; NAME/KEY: CDS  
; LOCATION: (42)-(2231)  
; NAME/KEY: modified\_base  
; LOCATION: (2270)  
; OTHER INFORMATION: n = g, a, c or t  
; OTHER INFORMATION: n = g, a, c or t  
US-09-287-849-1

## Alignment Scores:

Pred. No.: 9,78e-65 Length: 2287  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-287-849-1 (1-2287)

Qy 1 ThrAlaSerAspAsnPhelLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 63 ACGGCGCGTCCGATACTTCCAGCTGCCAGGGTGGGAGGATTCGCCATTCGGATC 132  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 123 GGGCAGCGATCGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTCAT 182  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 183 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACACAAACGCGCAACGGCGCACGA 242  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 243 GTCCACCGTGTTCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCGCTG 302  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 303 ATCAGCGGTCGACGGGCTCCGATCACTCGGCCACCGGATGGCGGACGGCGCTTAAC 362  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 363 GGGCATCATCCGGTACGTATCTCGGTGACCTTGGGCAACCAAGTCGGCGGCGACCGGT 422  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 423 ACAGGACGTGACATTTGGCGGAGGACCCCGGCC 458

## RESULT 4

US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.

; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

## Alignment Scores:

Pred. No.: 2,03e-60 Length: 4403765  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-103-840A-2 (1-4403765)

Qy 1 ThrAlaSerAspAsnPhelLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 151984 ACGGCGCGTCCGATACTTCCAGCTGCCAGGGTGGGAGGATTCGCCATTCGGATC 152043  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 152044 GGGCAGCGATCGCGATCGCGGCCAGATCCGATCGGGTGGGGGTACCCACCGTTCAT 152103  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 152104 ATCGGGCTACCGCTTCTCGGCTTGGTGTTCGACACACAGCGCAACGGCGCACGA 152163  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 152164 GTCCACCGTGTTCGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGCACGTG 152223  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 152224 ATCAGCGGTCGACGGGCTCCGATCACTCGGCCACCGGATGGCGGACGGCGCTTAAC 152283  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyThrArg 120  
Db 152284 GGGCATCATCCGGTACGTATCTCGGTGACCTTGGCAACCAAGTCGGCGGCGACCGGT 152343  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 152344 ACAGGACGTGACATTTGGCGGAGGACCCCGGCC 152379

## RESULT 5

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1

LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37RV  
US-09-103-840A-1  
Alignment Scores:  
Pred. No.: 2,04e-60 Length: 4411529  
Score: 675.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-09-684-215B-4 (1-132) x US-09-103-840A-1 (1-4411529)  
Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 151815 ACGGCGCGTCCGATTAATTCAGCTGTCCAGGTGGGAGGATTCGGCATTCGGATC 151874  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 151875 GGGCAGCGGATGGCGATCGGGCCAGATCCGATCGGTGGGGGTCACCACCGTTTCAT 151934  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60  
Db 151935 ATCGGGCTTACCGCTTCTCGCTTGGTGTGTTCGACAAACAGCGCAACGGCGCACGA 151994  
Qy 61 ValGlnArgValValGlySerAlaProAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 151995 GTCCAAACGGTGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGCGTG 152054  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 152055 ATCAACGGGTGTCGGCGCTCCGATCACTCGGCACCGGATGGCGGCGCGCTTAAAC 152114  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 152115 GGGCATCATCCCGGTGAGTGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 152174  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 152175 ACAGGGAACGTGACATTGGCCGAGGAGCCCCCGGCC 152210  
RESULT 6  
US-08-818-112-4  
Sequence 4, Application US/08818112  
Patent No. 6290969  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 153  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818.112

FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-818-112-4  
Alignment Scores:  
Pred. No.: 4,05e-65 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 3 Gaps: 0  
US-09-684-215B-4 (1-132) x US-08-818-112-4 (1-447)  
Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGGCGCGTCCGATTAATTCAGCTGTCCAGGTGGGAGGATTCGGCATTCGGATC 70  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 71 GGGCAGCGGATGGCGATCGGGCCAGATCCGATCGGTGGGGGTCACCACCGTTTCAT 130  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60  
Db 131 ATCGGGCTTACCGCTTCTCGCTTGGTGTGTTCGACAAACAGCGCAACGGCGCACGA 190  
Qy 61 ValGlnArgValValGlySerAlaProAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCAAACGGTGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCAACGGGTGTCGGCGCTCCGATCACTCGGCACCGGATGGCGGCGCGCTTAAAC 310  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCCGGTGAGTGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 370  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 371 ACAGGGAACGTGACATTGGCCGAGGAGCCCCCGGCC 406  
RESULT 7  
US-08-818-111-4  
Sequence 4, Application US/08818111  
Patent No. 6338552  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle

STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/818,111  
 FILING DATE: 13-MAR-1997  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.417C6  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 447 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-818-111-4

Alignment Scores:  
 Pred. No.: 4,05e-65 Length: 447  
 Score: 670.00 Matches: 131  
 Percent Similarity: 99.24% Conservative: 0  
 Best Local Similarity: 99.24% Mismatches: 1  
 Query Match: 99.26% Indels: 0  
 DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-08-818-111-4 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 11 ACGGCCGGTCCGATTAATTCAGCTGCCAGGGTGGCGAGGATTCCGCAATCCGATC 70  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 71 GGGCAGGGATGGCGATTCGGGGCCAGATCCGATCGGGTGGGGGTGACCCACCGTTTAT 130  
 Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
 Db 131 ATCCGGCTTACCGCTTCTCGCTTGGGTGTGTTCGACAAACACCGCAACGGCGCAGCA 190  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 191 GTCCACCGGTGGTCCGGAGCCCTCCGGCGGCAAGTCTCGGCATCTCCACGGCGACGTG 250  
 Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 251 ATCCCGCGGTGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 310  
 Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 Db 311 GGGCATCATCCCGGTGACGTATCTCGGTGAACCTGGCAAACTGGAGTCCGGCGGCGACGGT 370  
 Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 371 ACAGGGAACGTGACATTGGCCCGGAGGACCCCGGCC 406

RESULT 8

US-09-056-556-4  
 ; Sequence 4, Application US/09056556  
 ; Patent No. 6350456  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, Davin C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
 NUMBER OF SEQUENCES: 241  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED AND BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/056,556  
 FILING DATE: 07-APR-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.457  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 447 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-056-556-4

Alignment Scores:  
 Pred. No.: 4,05e-65 Length: 447  
 Score: 670.00 Matches: 131  
 Percent Similarity: 99.24% Conservative: 0  
 Best Local Similarity: 99.24% Mismatches: 1  
 Query Match: 99.26% Indels: 0  
 DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-056-556-4 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 Db 11 ACGGCCGGTCCGATTAATTCAGCTGCCAGGGTGGCGAGGATTCCGCAATCCGATC 70  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
 Db 71 GGGCAGGGATGGCGATTCGGGGCCAGATCCGATCGGGTGGGGGTGACCCACCGTTTAT 130  
 Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAlaArg 60  
 Db 131 ATCCGGCTTACCGCTTCTCGCTTGGGTGTGTTCGACAAACACCGCAACGGCGCAGCA 190  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 191 GTCCACCGGTGGTCCGGAGCCCTCCGGCGGCAAGTCTCGGCATCTCCACGGCGACGTG 250  
 Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 251 ATCCCGCGGTGACGGGCTCCGATCAACTCGGCCACCGGATGGCGGACGCGCTTAAC 310  
 Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 Db 311 GGGCATCATCCCGGTGACGTATCTCGGTGAACCTGGCAAACTGGAGTCCGGCGGCGACGGT 370  
 Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
 Db 371 ACAGGGAACGTGACATTGGCCCGGAGGACCCCGGCC 406

RESULT 9

US-09-072-596-4

TREAT

Sequence 4, Application US/09072596  
Patent No. 6458366  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-596-4

Alignment Scores:  
Pred. No.: 4,05e-65 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-072-596-4 (1-447)

QY	1	ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	11	ACGGCCGGTCCGATACCTTCAGCTGTCAGGGTGGCAGGGATTCGCCATTCGGATC	70
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40
Db	71	GGGCAGGCGATGCGATCGGGGCCAGATCCGATCGGTGGGGGTCAACCCCGTTCAT	130
QY	41	IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg	60
Db	131	ATCGGGCTACCCGCTTCCTCGGCTTGGGTGTCTGCACACACACGCGCAGCGACA	190
QY	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	191	GTCCACGCGTGTGCGGAGCGTCCGGCGCAAGTCTCGGATCTCCACCGCGCAGGTG	250
QY	81	IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAlaLeuAsn	100

Db 251 ATCACCGCGTGCAGCGGCTCCGATCACTCGGCCACCGCGATGGCGGACGGCTTAAC 310  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCCGGTGACGTCTATCTCGGTGAACCAACCAAGTCGGGCGGCGCGGT 370  
QY 121 ThrGlyAsnValThrLeuAlaGlyGlyProProAla 132  
Db 371 ACAGGGAACGTGACATTGGCCGAGGAGGCCCGCGCC 406

RESULT 10  
US-09-072-967-4  
Sequence 4, Application US/09072967  
Patent No. 6592877  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendrickson, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-967-4

Alignment Scores:  
Pred. No.: 4,05e-65 Length: 447  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-072-967-4 (1-447)

QY	1	ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle	20
Db	11	ACGGCCGGTCCGATACCTTCAGCTGTCAGGGTGGCAGGGATTCGCCATTCGGATC	70
QY	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis	40

Db 71 GGGCAGCGGATCGGATCGCGGCGGAGATCCGATCGGGTGGGGGTCAACCCACCGTTTCAT 130  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaSerAsnGlyAsnGlyValArg 60  
Db 131 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTCGACAAACGCGACGCGGACGCA 190  
QY 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyLeuGlyValArg 80  
Db 191 GTCCACGCGGTGCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGCGTG 250  
QY 81 IleThrAlaValAlaAspGlyAlaProLeuAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGCGGTGACGCGGCTCCGATCACTCGGCGACCGGATGCGGACGCGCTTAAC 310  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCGCGTACGTCATCTCGGTGAACCTGCAACCAAGTCGGGCGGCGACGCGT 370  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 371 ACAGGGAACGTGACATTCGCGGAGGACCCCGCGCC 406

## RESULT 11

US-08-818-112-17  
; Sequence 17, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Hedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1872 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-818-112-17

; Sequence 17, Application US/08818112

; Patent No. 6290969

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Houghton, Raymond

; APPLICANT: Hedvick, Thomas S.

; APPLICANT: Twardzik, Daniel R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS

; NUMBER OF SEQUENCES: 153

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/818,112

; FILING DATE: 13-MAR-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.411C6

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1872 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-818-112-17

Alignment Scores:

Pred. No.: 2,66e-64

Score: 670.00

Percent Similarity: 99.24%

Length: 1872

Matches: 131

Conservative: 0

Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 3 Gaps: 0

US-09-684-215B-4 (1-132) x US-08-818-112-17 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGGCGCGGTCGATACCTTCAGCTGTCCAGGGTGGCAGGATTCGCCATTCCGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 818 GGGCAGGCGATGGCGATCCGGGCGAAATCCGATCGGGTGGGGGTCCACCACCGTTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaSerAsnGlyGlyValArg 60  
Db 878 ATCGGCGCTACCGCTTCCTCGGCTTGGGTGTTCGACAAACACGCGCAACGCGCACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCCAACGCGTGGTCCGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGAGTG 997  
QY 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCACCGCGGTGACGCGGCTCCGATCACTCGGCGACCGCGATGGCGGACGCGCTTAAC 1057  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 1058 GGGCATCATCCGCGTACGTCATCTCGGTGAACCTGCAACCAAGTCGGGCGGCGACGCGT 1117  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 1118 ACAGGGAACGTGACATTCGCGGAGGACCCCGCGCC 1153

## RESULT 12

US-08-818-111-17  
; Sequence 17, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Campos-Neto, Antonio

; APPLICANT: Houghton, Raymond

; APPLICANT: Hedvick, Thomas S.

; APPLICANT: Twardzik, Daniel R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF

; NUMBER OF SEQUENCES: 148

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/818,111

; FILING DATE: 13-MAR-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.417C6

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:



```

; LENGTH: 1872 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-818-111-17

Alignment Scores:
Pred. No.: 2,66e-64 Length: 1872
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-08-818-111-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 758 ACGGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGGCGAGGATTCGCCATTCGCATC 817
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 818 GGGCAGGCGATCGGCTTCCTCGCTTGGGTGTGTGACAAACACGGCAACGGCGCACGA 937
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 938 GTCCAAACGGTGGTGGAGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 997
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 998 ATCACCAGGCGTCCAGCGCTCCGATCAACTCGGCGCACGGGATGGCGAGCGCGTTAAC 1057
Qy 101 GlyHisHisProGlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArg 120
Db 1058 GGGCATCATCCCGGTGACGTATCTCGGTGAACCTGGCAACCAAGTCGGGGGCGACGCGT 1117
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 1118 ACAGGGAACGTGACATTTGGCCCGAGGACCCCGGGCC 1153

RESULT 13
US-09-056-556-17
; Sequence 17, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392

TREATM

; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-056-556-17

Alignment Scores:
Pred. No.: 2,66e-64 Length: 1872
Score: 670.00 Matches: 131
Percent Similarity: 99.24% Conservative: 0
Best Local Similarity: 99.24% Mismatches: 1
Query Match: 99.26% Indels: 0
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-056-556-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 758 ACGGCGCGTCCGATTAACCTCCAGCTGCCAGGGTGGGCGAGGATTCGCCATTCGCATC 817
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40
Db 818 GGGCAGGCGATCGGCTTCCTCGCTTGGGTGTGTGACAAACACGGCAACGGCGCACGA 937
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60
Db 878 ATCGGCGCTACCGCTTCCTCGCTTGGGTGTGTGACAAACACGGCAACGGCGCACGA 937
Qy 878 ATCGGCGCTACCGCTTCCTCGCTTGGGTGTGTGACAAACACGGCAACGGCGCACGA 937
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 938 GTCCAAACGGTGGTGGAGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 997
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 998 ATCACCAGGCGTCCAGCGCTCCGATCAACTCGGCGCACGGGATGGCGAGCGCGTTAAC 1057
Qy 101 GlyHisHisProGlyAspValIleSerValThrTpGlnThrLysSerGlyGlyThrArg 120
Db 1058 GGGCATCATCCCGGTGACGTATCTCGGTGAACCTGGCAACCAAGTCGGGGGCGACGCGT 1117
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
Db 1118 ACAGGGAACGTGACATTTGGCCCGAGGACCCCGGGCC 1153

RESULT 14
US-09-072-596-17
; Sequence 17, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvik, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; TUBERCULOSIS
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,596  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.417C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-596-17

Alignment Scores:  
Pred. No.: 2,66e-64 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-072-596-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGCCGCGTCCGATTAATCCAGCTGTCCAGGCTGGGCGGAGGATTCGCCATTCGATC 817  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 818 GGCAGCGGATGCGGATCCGAGGCGAAATCCGATCGGTGGGGGTCAACCGCTTCAAT 877  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArg 60  
Db 878 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGCAACAACGCGGCGGCGACGA 937  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCACACGCTGTCGAGACGCTCCGCGGCGCAAGTCTCGGATCTCCACCGCGACGTG 997  
Qy 81 IleThrAlaValaspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCAGCGGCTGACGCGCTCCGATCAACTCGGCCACCGCGATGCGGCGCGCTTAAC 1057  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 1058 GGCATCATCCCGGTGACGTCATCTCGGTGAACCTGGCAACCAAGTCCGGGCGGCGCGGT 1117  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 1118 ACAGGGAACGTGACATTTGGCGGAGGACCCCGGCC 1153

RESULT 15

US-09-072-967-17  
Sequence 17, Application US/09072967  
Patent No. 6592877  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedwick, Thomas S.

APPLICANT: Twardzik, Daniel R.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Hendricksen, Ronald C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 355  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,967  
FILING DATE: 05-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.411C9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-072-967-17

Alignment Scores:  
Pred. No.: 2,66e-64 Length: 1872  
Score: 670.00 Matches: 131  
Percent Similarity: 99.24% Conservative: 0  
Best Local Similarity: 99.24% Mismatches: 1  
Query Match: 99.26% Indels: 0  
DB: 4 Gaps: 0

US-09-684-215B-4 (1-132) x US-09-072-967-17 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGCCGCGTCCGATTAATCCAGCTGTCCAGGCTGGGCGGAGGATTCGCCATTCGATC 817  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 818 GGCAGCGGATGCGGATCCGAGGCGAAATCCGATCGGTGGGGGTCAACCGCTTCAAT 877  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArg 60  
Db 878 ATCGGCGCTACCGCTTCTCGGCTTGGGTGTGTGCAACAACGCGGCGGCGACGA 937  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCACACGCTGTCGAGACGCTCCGCGGCGCAAGTCTCGGATCTCCACCGCGACGTG 997  
Qy 81 IleThrAlaValaspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCAGCGGCTGACGCGCTCCGATCAACTCGGCCACCGCGATGCGGCGCGCTTAAC 1057  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 1058 GGCATCATCCCGGTGACGTCATCTCGGTGAACCTGGCAACCAAGTCCGGGCGGCGCGGT 1117  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Mon May 3 06:36:09 2004

Db 1118 ACAGGGAACGTGACATTGGCCGAGGAGGACCCCGGCC 1153

## RESULT 16

US-09-636-215-822

; Sequence 822, Application US/09636215

; Patent No. 6620922

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.42717C17

; CURRENT APPLICATION NUMBER: US/09/636,215

; CURRENT FILING DATE: 2000-08-10

; NUMBER OF SEQ ID NOS: 852

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-636-215-822

Alignment Scores:

Pred. No.:	1.03e-60	Length:	675
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-636-215-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 22 ACGCCCGGTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGGATTCCGATTCGGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

Db 82 GGGCAGCGGTGGCGATCGCGGGCCAGATCAAG-----CTTCCACCGTTTCAT 129

Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60

Db 130 ATCGGGCTTACCCCTTCCTCGCTTGGTGTGTGTCGACAAACACGGCAACGGCGCACGA 189

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 190 GTCCAAACGGGTGTCGGAGCGTCCCGGGGCAAGTCTCGGATCTCCACCGCGACGTG 249

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 250 ATCACCAGGTTCGACGGCGTCCGATCACTCGGCCACCGCGATGCGACGGCTTAAC 309

Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg 120

Db 310 GGGCATCATCCCGTGACGTCATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 369

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 370 ACAGGGAACGTGACATTGGCCGAGGAGGACCCCGGCC 405

## RESULT 17

US-09-685-166A-822

; Sequence 822, Application US/09685166A

; Patent No. 6630305

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C21

; CURRENT APPLICATION NUMBER: US/09/685,166A

; CURRENT FILING DATE: 2000-10-10

; NUMBER OF SEQ ID NOS: 898

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 822

; LENGTH: 675

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-685-166A-822

Alignment Scores:

Pred. No.:	1.03e-60	Length:	675
Score:	632.00 <td>Matches:</td> <td>126</td>	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-685-166A-822 (1-675)

Qy 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20

Db 22 ACGCCCGGTCCGATTAACCTCCAGCTGTCCAGGGTGGCAGGGATTCCGATTCGGATC 81

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40

Db 82 GGGCAGCGGTGGCGATCGCGGGCCAGATCAAG-----CTTCCACCGTTTCAT 129

Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValAlaArg 60

Db 130 ATCGGGCTTACCCCTTCCTCGCTTGGTGTGTGTCGACAAACACGGCAACGGCGCACGA 189

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 190 GTCCAAACGGGTGTCGGAGCGTCCCGGGGCAAGTCTCGGATCTCCACCGCGACGTG 249

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 250 ATCACCAGGTTCGACGGCGTCCGATCACTCGGCCACCGCGATGCGACGGCTTAAC 309

Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlyssSerGlyGlyThrArg 120

Db 310 GGGCATCATCCCGTGACGTCATCTCGTGACCTGGCAACCAAGTCCGGCGGCGACGCGT 369

Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 370 ACAGGACGTGACATTGGCCGAGGAGCCCGCGCC 405

## RESULT 18

US-09-736-457-1862  
; Sequence 1862, Application US/09736457

; Patent No. 6509448

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedwick, Tom

; APPLICANT: Carter, Darick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Aijun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457

; CURRENT FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 1864

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1862

; LENGTH: 822

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-736-457-1862

## Alignment Scores:

Pred. No.: 1,346-60 Length: 822  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 4 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-736-457-1862 (1-822)

Qy 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACAGGCGGTCCGATTAATCCAGTGTCCAGGTTGGCGGAGGATTCGCCATTCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGGTATGGCGATCGGGCCAGATCAAG-----CTTCCACCGTTTAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAlaSerAsnGlyAenGlyAlaArg 60  
Db 130 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGACGCGACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCACGCGTGTGGGAGCGTCCGCGCGCAAGTCTCCGATCTCCACCGCGACGCG 249  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGCGTGCAGCGCGTCCGATCACTCGGCCACCGCGATGGCGAGCGCTTAAC 309  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyThrArg 120  
Db 310 GGGCATCATCCCGTGCAGTCACTCGGTGACCTGGCAACCAAGTCGGCGCGCGCGT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGACGTGACATTGGCCGAGGAGCCCGCGCC 405

## RESULT 19

US-09-643-597-353

; Sequence 353, Application US/09643597

; Patent No. 6426072

; GENERAL INFORMATION:

; APPLICANT: Fanger, Gary R.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Wang, Liqun

; APPLICANT: Fanger, Gary R.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Wang, Liqun

; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 353  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-643-597-353

Alignment Scores:  
Pred. No.: 1,516-60 Length: 900  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: 4 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-643-597-353 (1-900)

Qy 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 22 ACAGGCGGTCCGATTAATCCAGTGTCCAGGTTGGCGGAGGATTCGCCATTCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGGTATGGCGATCGGGCCAGATCAAG-----CTTCCACCGTTTAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyValValAlaSerAsnGlyAenGlyAlaArg 60  
Db 130 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTGTCGACAAACGCGACGCGACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCACGCGTGTGGGAGCGTCCGCGCGCAAGTCTCCGATCTCCACCGCGACGCG 249  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGCGTGCAGCGCGTCCGATCACTCGGCCACCGCGATGGCGAGCGCTTAAC 309  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyThrArg 120  
Db 310 GGGCATCATCCCGTGCAGTCACTCGGTGACCTGGCAACCAAGTCGGCGCGCGCGT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGACGTGACATTGGCCGAGGAGCCCGCGCC 405

RESULT 20  
US-09-608-421B-353  
; Sequence 353, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liqun  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.

APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.455C9  
CURRENT APPLICATION NUMBER: US/09/606,421B  
CURRENT FILING DATE: 2000-06-28  
NUMBER OF SEQ ID NOS: 358  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 353  
LENGTH: 900  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-606-421B-353

Alignment Scores: 1,518-60 Length: 900  
Pred. No.: 632.00 Matches: 126  
Score: 96.21% Conservative: 1  
Percent Similarity: 95.45% Mismatches: 1  
Best Local Similarity: 93.63% Indels: 4  
Query Match: 4 Gaps: 1  
DB:

US-09-684-215B-4 (1-132) x US-09-606-421B-353 (1-900)

Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGCGAGGATTCGCCATCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleAgsSerGlyGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGATGGCGATCCGGGCGGCGAGATCAAG-----CTTCCACCGCTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 130 ATCCGCGCTACCGCTTCCTCGCTTGGGTGTGTGCGAACAAACGCGACGCGACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCAAACGGGTGTCGGGAGCGCTCCGGCGGCGAGTCTCGGATCTCCACCGCGCGGTG 249  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCCACCGCGTCCGACGCGCTCCGATCAACTCGGCGACCGCGATGCGCGCGCTTAAC 309  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGTGACGTCATCTCGTGACCTGGCAACCAAGTCCGGCGCGCGCT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 405

RESULT 21  
US-09-636-215-834  
Sequence 834, Application US/09636215  
Patent No. 6620922  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: COMPOSITIONS OF PROSTATE CANCER  
FILE REFERENCE: 210121.42717C17  
CURRENT APPLICATION NUMBER: US/09/636,215  
CURRENT FILING DATE: 2000-08-10  
NUMBER OF SEQ ID NOS: 852  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 834  
LENGTH: 915  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-636-215-834

Alignment Scores: 1,548-60 Length: 915  
Pred. No.: 632.00 Matches: 126  
Score: 96.21% Conservative: 1  
Percent Similarity: 95.45% Mismatches: 1  
Best Local Similarity: 93.63% Indels: 4  
Query Match: 4 Gaps: 1  
DB:

US-09-684-215B-4 (1-132) x US-09-636-215-834 (1-915)

Qy 1 ThrAlaAlaSerAspAsnPhelGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGGCGCGTCCGATAAATTCAGCTGTCCAGGGTGGCGAGGATTCGCCATCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleAgsSerGlyGlyGlySerProThrValHis 40  
Db 82 GGGCAGCGATGGCGATCCGGGCGGCGAGATCAAG-----CTTCCACCGCTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 130 ATCCGCGCTACCGCTTCCTCGCTTGGGTGTGTGCGAACAAACGCGACGCGACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCAAACGGGTGTCGGGAGCGCTCCGGCGGCGAGTCTCGGATCTCCACCGCGCGGTG 249  
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCCACCGCGTCCGACGCGCTCCGATCAACTCGGCGACCGCGATGCGCGCGCTTAAC 309  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGTGACGTCATCTCGTGACCTGGCAACCAAGTCCGGCGCGCGCT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132  
Db 370 ACAGGGAACGTGACATTGGCGGAGGACCCCGGCC 405

RESULT 22  
US-09-685-166A-834  
Sequence 834, Application US/09685166A  
Patent No. 6630395  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C21  
; CURRENT APPLICATION NUMBER: US/09/685,166A  
; CURRENT FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 898  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 834  
; LENGTH: 915  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-685-166A-834

Alignment Scores:  
Pred. No.: 1,54e-60 Length: 915  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: Gaps: 1

US-09-684-215B-4 (1-132) x US-09-685-166A-834 (1-915)

Qy 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGCCCGCTCGGATACCTTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyValHis 40  
Db 82 GGGCAGCGGATGGCGATCGGGCCAGATCAAG-----CTTCCACCGTTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValArg 60  
Db 130 ATCCGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACACGCGCACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyVal 80  
Db 190 GTCCACGCGTGGTTCGGAGCGCTCCGGCGGAGTCTCGGCATCTCCACCGCGACGTG 249  
Qy 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeu 100  
Db 250 ATCCGCGCTCGACGCGCTCCGATCACTCGGCACCGCGATGGCGGACGCGTTAAC 309  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyValArg 120  
Db 310 GGGCATCATCCCGTGACGTATCTCGTGACCTGGCAACCAACAGTCCGGCGGACGCGT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 405

## RESULT 23

US-09-736-457-1861  
; Sequence 1861, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Restek, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457

; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1861  
; LENGTH: 945  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1861

Alignment Scores:  
Pred. No.: 1,61e-60 Length: 945  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1  
Query Match: 93.63% Indels: 4  
DB: Gaps: 1

US-09-684-215B-4 (1-132) x US-09-736-457-1861 (1-945)

Qy 1 ThrAlaAlaSerAspAsnPhgInLeuSerGlnGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACGCCCGCTCGGATACCTTCCAGCTGTCAGGCTGGGAGGATTCGCCATTCCGATC 81  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyValHis 40  
Db 82 GGGCAGCGGATGGCGATCGGGCCAGATCAAG-----CTTCCACCGTTTCAT 129  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyValArg 60  
Db 130 ATCCGCGCTACCGCTTCTCCGCTTGGGTGTTGTCGACAAACACGCGCACGA 189  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyVal 80  
Db 190 GTCCACGCGTGGTTCGGAGCGCTCCGGCGGAGTCTCGGCATCTCCACCGCGACGTG 249  
Qy 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeu 100  
Db 250 ATCCGCGCTCGACGCGCTCCGATCACTCGGCACCGCGATGGCGGACGCGTTAAC 309  
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyValArg 120  
Db 310 GGGCATCATCCCGTGACGTATCTCGTGACCTGGCAACCAACAGTCCGGCGGACGCGT 369  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 405

## RESULT 24

US-09-643-597-351  
; Sequence 351, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 351  
; LENGTH: 1012  
; TYPE: DNA

ORGANISM: Homo sapiens  
US-09-643-597-351

Alignment Scores:  
Pred No.: 1,76e-60 Length: 1012  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 4  
Query Match: 93.63% Indels: 4  
DB: Gaps: 1

US-09-684-215B-4 (1-132) x US-09-643-597-351 (1-1012)

Qy 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 67 ACGCCCGCTCCGATACTTCCAGCTGTCACGGGTGGGAGGATTCGCCATCCGATC 126  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 127 GGGCAGCGGATGCGGATCGCGGCGAGATCAAG-----CTTCCACCGTTTCAAT 174  
Qy 41 IledGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
Db 175 ATCCGGGCTTCCGCTTCCGCTTGGGTGTGTCGACAAACGCGGACGCGACGA 234  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 235 GTCCAAACGCGTGTGGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 294  
Qy 81 IledThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 295 ATCCCGCGGTGACGCGCTCCGATCACTCGGCACCGGATGGCGGCGGTTCATAC 354  
Qy 101 GlyHisProGlyAspValIleSerValThrTTPGlnThrLysSerGlyGlyThrArg 120  
Db 355 GGGCATCATCCGCTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGT 414  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 415 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 450

RESULT 25

US-09-606-421B-351  
Sequence 351, Application US/09606421B  
Patent No. 6511315

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong  
APPLICANT: Fan, Liqun  
APPLICANT: Kalos, Michael D.  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Hosken, Nancy  
APPLICANT: Fanger, Gary R.  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.455C9  
CURRENT APPLICATION NUMBER: US/09/606,421B  
CURRENT FILING DATE: 2000-06-28  
NUMBER OF SEQ ID NOS: 358  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 351  
LENGTH: 1012  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-606-421B-351

Alignment Scores:  
Pred. No.: 1,76e-60 Length: 1012  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1

Query Match: 93.63% Indels: 4  
DB: Gaps: 1

US-09-684-215B-4 (1-132) x US-09-606-421B-351 (1-1012)

Qy 1 ThrAlaAlaSerAspAsnPhenLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 67 ACGCCCGCTCCGATACTTCCAGCTGTCACGGGTGGGAGGATTCGCCATCCGATC 126  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 127 GGGCAGCGGATGCGGATCGCGGCGAGATCAAG-----CTTCCACCGTTTCAAT 174  
Qy 41 IledGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyAsnGlyAlaArg 60  
Db 175 ATCCGGGCTTCCGCTTCCGCTTGGGTGTGTCGACAAACGCGGACGCGACGA 234  
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 235 GTCCAAACGCGTGTGGGAGCGTCCGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 294  
Qy 81 IledThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 295 ATCCCGCGGTGACGCGCTCCGATCACTCGGCACCGGATGGCGGCGGTTCATAC 354  
Qy 101 GlyHisProGlyAspValIleSerValThrTTPGlnThrLysSerGlyGlyThrArg 120  
Db 355 GGGCATCATCCGCTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGGCGGCGGT 414  
Qy 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 415 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 450

RESULT 26

US-09-636-215-851  
Sequence 851, Application US/09636215  
Patent No. 6620922

GENERAL INFORMATION:

APPLICANT: Xu, Jianshun  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darzick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.42717C17  
CURRENT APPLICATION NUMBER: US/09/636,215  
CURRENT FILING DATE: 2000-08-10  
NUMBER OF SEQ ID NOS: 852  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 851  
LENGTH: 1203  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-636-215-851

Alignment Scores:  
Pred. No.: 2,21e-60 Length: 1203  
Score: 632.00 Matches: 126  
Percent Similarity: 96.21% Conservative: 1  
Best Local Similarity: 95.45% Mismatches: 1

```

Query Match: 93.63% Indels: 4
DB: 4 Gaps: 1
US-09-684-215B-4 (1-132) x US-09-636-215-851 (1-1203)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGCGATGGCGATCGCGGCGAGATCAAG-----CTTCCCACCGTTTCAT 129
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 130 ATCGGGCTTACCGCTTCCCTGGCTTGGGTGTGTGTCGACAAACGCGGACGGCGACGA 189
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCAAACGGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 249
Qy 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATACCGCGGTGGACGGCGCTCCGATCACTCCGCGCACCGGATGGCGGACGGCTTAAAC 309
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGGT 369
Qy 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 370 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 405

RESULT 27
US-09-685-166A-851
; Sequence 851, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-851

Alignment Scores:
Pred. No.: 2,216-60 Length: 1203
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1

US-09-684-215B-4 (1-132) x US-09-685-166A-851 (1-1203)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGCGATGGCGATCGCGGCGAGATCAAG-----CTTCCCACCGTTTCAT 129
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 130 ATCGGGCTTACCGCTTCCCTGGCTTGGGTGTGTGTCGACAAACGCGGACGGCGACGA 189
Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 190 GTCCAAACGGGTGGTGGGAGCGCTCCGGCGCAAGTCTCGGCATCTCCACCGGCGACGTG 249
Qy 81 IleThrAlaValAlaAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 250 ATACCGCGGTGGACGGCGCTCCGATCACTCCGCGCACCGGATGGCGGACGGCTTAAAC 309
Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 310 GGGCATCATCCGGTGACGTCACTCGGTGACCTGGCAACCAAGTCCGGCGGCGACGGT 369
Qy 121 ThrGlyAsnValThrLeuAlaGlyProProAla 132
Db 370 ACAGGAACGTGACATTGGCGGAGGACCCCGGCC 405

RESULT 28
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6449234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Alignment Scores:
Pred. No.: 2,866-60 Length: 1464
Score: 632.00 Matches: 126
Percent Similarity: 96.21% Conservative: 1
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 93.63% Indels: 4
DB: 4 Gaps: 1

US-09-684-215B-4 (1-132) x US-09-620-412C-348 (1-1464)

Qy 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20
Db 22 ACGCCCGCTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 81
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 82 GGGCAGCGATGGCGATCGCGGCGAGATCAAG-----CTTCCCACCGTTTCAT 129
Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 130 ATCGGGCTTACCGCTTCCCTGGCTTGGGTGTGTGTCGACAAACGCGGACGGCGACGA 189

```



QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCACCGCGTGGGAGCGCTCCCGCGCGAAGTCTCGGCATCTCCACCGCGACG 249  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGCGTGGAGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGTTAAC 309  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGTGGAGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGTTAAC 369  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGAAACGTGACATTGGCGGAGGACCCCGGCC 405

## RESULT 29

US-09-598-419-348

; Sequence 348, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C5

; CURRENT APPLICATION NUMBER: US/09/598,419

; CURRENT FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 357

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 348

; LENGTH: 1464

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-598-419-348

## Alignment Scores:

Pred. No.:	2,86e-60	Length:	1464
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-598-419-348 (1-1464)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACAGCGCGTCCGATCAACTCCAGCTGTCACAGGTTGGCGAGGATTCGCCATTCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGGCGCATGGCGATCCCGGCGCATCAAG-----CTTCCACCGTTTAT 129  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 130 ATCCGGGCTACCGGCTTCTCCGCTGGGTGTTGTCACAAACGCGCACGCGCACGA 189  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCACCGGTGGTGGAGCGCTCCGCGGCGCATCTCCGATCTCCACCGCGACG 249  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGCGTGGAGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGTTAAC 309  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGTGGAGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGTTAAC 369  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db 370 ACAGGAAACGTGACATTGGCGGAGGACCCCGGCC 405

## RESULT 30

US-09-620-412C-332

; Sequence 332, Application US/09620412C

; Patent No. 6448234

; GENERAL INFORMATION:

; APPLICANT: Steven P. Fling

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C7

; CURRENT APPLICATION NUMBER: US/09/620,412C

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 332

; LENGTH: 1557

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-620-412C-332

## Alignment Scores:

Pred. No.:	3.1e-60	Length:	1557
Score:	632.00	Matches:	126
Percent Similarity:	96.21%	Conservative:	1
Best Local Similarity:	95.45%	Mismatches:	1
Query Match:	93.63%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-4 (1-132) x US-09-620-412C-332 (1-1557)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 22 ACAGCGCGTCCGATCAACTCCAGCTGTCACAGGTTGGCGAGGATTCGCCATTCGATC 81  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
Db 82 GGGCAGGCGATGGCGATCCCGGCGCATCAAG-----CTTCCACCGTTTAT 129  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 130 ATCCGGGCTACCGGCTTCTCCGCTGGGTGTTGTCACAAACGCGCACGCGCACGA 189  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 190 GTCCACCGGTGGTGGAGCGCTCCGCGGCGCATCTCCGATCTCCACCGCGACG 249  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 250 ATCACCAGCGTGGAGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGTTAAC 309  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 310 GGGCATCATCCCGTGGAGCGCTCCGATCAACTCGGCCACCGCATGGCGGACGCGTTAAC 369  
QY 121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132  
Db 370 ACAGGAAACGTGACATTGGCGGAGGACCCCGGCC 405

Search completed: April 30, 2004, 05:01:06

Job time : 1085.58 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 29, 2004, 20:32:52 ; Search time 3670.05 Seconds

(without alignments)  
1511.672 Million cell updates/sec

Title: US-09-684-215B-23

Perfect score: 651

Sequence: 1 TRAADNFQLSQGGGAIRI.....SVTWQTKSGGTRIGNVTLAE 128

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09684215/runat\_29042004\_061304\_13185/app\_query.fasta\_1.1180  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09684215@cgn 1.1 6034 @runat\_29042004\_061304\_13185 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.btg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	651	100.0	702	6	BD251334	Fused pro
2	651	100.0	702	6	AR403747	AR403747 Sequence
3	651	100.0	1068	6	BD274032	BD274032 Sequence
4	651	100.0	1068	6	AX005788	AX005788 Sequence
5	651	100.0	1143	6	BD274033	BD274033 Sequence
6	651	100.0	1143	6	AX005790	AX005790 Sequence
7	651	100.0	2287	6	BD251322	BD251322 Fused pro
8	651	100.0	2287	6	AR303127	AR303127 Sequence
9	651	100.0	2287	6	AR403735	AR403735 Sequence
10	651	100.0	14029	1	AR006925	AE006925 Mycobacte
11	651	100.0	341957	15	BX842572	Bx842572 Mycobacte
12	651	100.0	343050	1	BX248334	BX248334 Mycobacte
13	646	99.2	447	6	AR169152	AR169152 Sequence
14	646	99.2	447	6	AR182442	AR182442 Sequence
15	646	99.2	447	6	AR194825	AR194825 Sequence
16	646	99.2	447	6	AR233097	AR233097 Sequence
17	646	99.2	447	6	AR353302	AR353302 Sequence
18	646	99.2	447	6	AX429596	AX429596 Sequence
19	646	99.2	447	6	AX832581	AX832581 Sequence
20	646	99.2	447	6	BD006325	BD006325 Compounds
21	646	99.2	447	6	BD006445	BD006445 Compounds
22	646	99.2	447	6	BD069285	BD069285 Compounds
23	646	99.2	447	6	BD205817	BD205817 Compounds
24	646	99.2	1872	6	AR169165	AR169165 Sequence
25	646	99.2	1872	6	AR182455	AR182455 Sequence
26	646	99.2	1872	6	AR194838	AR194838 Sequence
27	646	99.2	1872	6	AR233110	AR233110 Sequence
28	646	99.2	1872	6	AR353315	AR353315 Sequence
29	646	99.2	1872	6	AX429609	AX429609 Sequence
30	646	99.2	1872	6	AX832594	AX832594 Sequence
31	646	99.2	1872	6	BD006338	BD006338 Compounds
32	646	99.2	1872	6	BD006458	BD006458 Compounds
33	646	99.2	1872	6	BD069298	BD069298 Compounds
34	646	99.2	1872	6	BD205830	BD205830 Compounds
35	608	93.4	675	6	AR261272	AR261272 Sequence
36	608	93.4	675	6	AR400535	AR400535 Sequence
37	608	93.4	675	6	AR405802	AR405802 Sequence
38	608	93.4	675	6	AX201049	AX201049 Sequence
39	608	93.4	675	6	AX267848	AX267848 Sequence
40	608	93.4	822	6	AR277645	AR277645 Sequence
41	608	93.4	822	6	AX369152	AX369152 Sequence
42	608	93.4	894	6	AX351489	AX351489 Sequence
43	608	93.4	900	6	AR220690	AR220690 Sequence
44	608	93.4	900	6	AX365960	AX365960 Sequence
45	608	93.4	915	6	AR261281	AR261281 Sequence

ALIGNMENTS

RESULT 1

BD251334  
 LOCUS 702 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Fused protein of Mycobacterium tuberculosis antigen and utilization thereof.  
 ACCESSION BD251334  
 VERSION BD251334.1 GI:33061104  
 KEYWORDS JP 2002510494-A/13.  
 SOURCE synthetic construct  
 ORGANISM artificial construct  
 REFERENCE 1 (bases 1 to 702)  
 AUTHORS Skeiky, Y.A.W., Alderson, M. and Neto, A.C.  
 TITLE Fused protein of Mycobacterium tuberculosis antigen and utilization thereof  
 JOURNAL Patent: JP 2002510494-A 13 09-APR-2002;  
 COMMENT CORIXA CORP  
 OS Artificial Sequence  
 FN JP 2002510494-A/13  
 PD 09-APR-2002  
 PF 07-APR-1999 JP 2000542460  
 PR 07-APR-1998 US 09/056556, 30-DEC-1998 US 09/223040 PI  
 VASIR A W SKEIKY, MARK ALDERSON, ANTONIO CAMPOS NETO PC  
 C12N15/09, A61K39/04, A61K48/00, A61P31/04, C07K14/35, C07K19/00, PC  
 C12P21/02  
 PC C12N15/00  
 CC Description of Artificial Sequence: bi-fusion  
 protein Ra12-DPPD  
 CC (designated  
 CC Mtb24), reading frame 1  
 CC bi-fusion protein Ra12-DPPD (designated Mtb24), reading frame 1  
 CC reading frame 2  
 CC reading frame 3  
 FH Key Location/Qualifiers  
 FT CDS (1)..(696)  
 FT CDS (2)..(700)  
 FT CDS (3)..(701).

FEATURES  
 source  
 1..702  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,55e-42 Length: 702  
 Score: 651.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x BD251334 (1-702)

Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProile 20  
 Db 25 ACGGCGCGTCCGATTAATTCAGCTGCCAGGTGCCAGGTGGCGAGGATTCGCCATTCGATC 84  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleAargSerGlyGlyGlySerProThrValHis 40  
 Db 85 GGGCAGCGATGGCGATCCGAGATCCGATCCGAGGTGGCGAGGATTCGCCATTCGATC 144  
 Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAsnAsnGlyValAlaArg 60  
 Db 145 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTCGACACACACGCGACGCGACGCA 204  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 205 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACGCGGACGTG 264  
 Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 265 ATCACCAGCGGTCCGAGCGGTCCGATCACTCACTCGGCCCGCGATGGCGGACGCTTAAC 324  
 Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGT 384  
 Qy 121 ThrGlyAsnValThrLeuAlaGlu 128  
 Db 385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 3  
 BD274032  
 LOCUS 1068 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Sequences nucleic acid from polypeptides exportes from mycobacteries, vector the complement and applications at diagnostic and the prevention from the tuberculosis.

Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGT 384  
 Qy 121 ThrGlyAsnValThrLeuAlaGlu 128  
 Db 385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 2  
 AR403747  
 LOCUS 702 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 27 from patent US 6627198.  
 ACCESSION AR403747  
 VERSION AR403747.1 GI:40151423  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 702)  
 AUTHORS Reed, S.G., Skeiky, Y.A., Dillon, D.C., Alderson, M. and Campos-Neto, A.  
 TITLE Fusion proteins of Mycobacterium tuberculosis antigens and their uses  
 JOURNAL Patent: US 6627198-A 27 30-SEP-2003;  
 FEATURES Location/Qualifiers  
 source  
 1..702  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,55e-42 Length: 702  
 Score: 651.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR403747 (1-702)

Qy 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
 Db 25 ACGGCGCGTCCGATTAATTCAGCTGCCAGGTGGCGAGGATTCGCCATTCGATC 84  
 Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleAargSerGlyGlyGlySerProThrValHis 40  
 Db 85 GGGCAGCGATGGCGATCCGAGATCCGATCCGAGGTGGCGAGGATTCGCCATTCGATC 144  
 Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAsnAsnGlyValAlaArg 60  
 Db 145 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTTCGACACACACGCGACGCGACGCA 204  
 Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 205 GTCCAAACGCGTGGTGGGAGCGCTCCGCGGCAAGTCTCGGCATCTCCACGCGGACGTG 264  
 Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 265 ATCACCAGCGGTCCGAGCGGTCCGATCACTCACTCGGCCCGCGATGGCGGACGCTTAAC 324  
 Qy 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
 Db 325 GGGCATCATCCCGGTGACGTATCTCGGTGACCTGGCAACCAAGTCCGGCGGACGCGT 384  
 Qy 121 ThrGlyAsnValThrLeuAlaGlu 128  
 Db 385 ACAGGGAACGTGACATTGGCCGAG 408

## RESULT 3

BD274032  
 LOCUS 1068 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Sequences nucleic acid from polypeptides exportes from mycobacteries, vector the complement and applications at diagnostic and the prevention from the tuberculosis.

ACCESSION BD274032  
 VERSION 1 GI:33083800  
 KEYWORDS JP 2002534956-A/256  
 SOURCE Mycobacterium tuberculosis  
 ORGANISM Mycobacterium tuberculosis  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
 tuberculosis complex.  
 REFERENCE 1 (bases 1 to 1068)  
 AUTHORS Gicquel, B., Portnoi, D., Lim, E., Pellicic, V., Guigueno, A. and  
 Salmoniere, Y.G.D.L.  
 TITLE Sequences nucleic acid from polypeptides exportes from  
 mycobacteria, vector the comprenant and applications at diagnostic  
 and the prevention from the tuberculose  
 JOURNAL Patent: JP 2002534956-A 256 22-OCT-2002;  
 INSTITUT PASTEUR  
 COMMENT OS Mycobacterium tuberculosis  
 PN JP 2002534956-A/256  
 PD 22-OCT-2002  
 PF 14-AUG-1998 JP 2000509849  
 PR 14-AUG-1997 FR 97/10404, 11-SEP-1997 FR 97/11325 PI  
 BRIGITTE GICQUEL, DENIS PORTNOI, ENG-WONG LIM, VLADIMIR PELICIC, PI  
 AGNES GUIGUENO, YVES GOGUET DE LA SALMONIERE  
 PC C12N15/09, A61K39/04, A61K48/00, A61P11/00, A61P31/04, C07K14/35,  
 PC C07K16/12,  
 PC C07K19/00, C12N1/21, C12P21/02, C12Q1/02, C12Q1/68, G01N33/50, PC  
 G01N33/53//  
 PC C12N15/09, C12R1/32, C12N15/00, C12N15/00, C12R1/32) CC 50D  
 PH Key Location/Qualifiers  
 FT CDS  
 FEATURES source  
 1..1068  
 /organism="Mycobacterium tuberculosis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:1773"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,428-42 Length: 1068  
 Score: 651.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x BD274032 (1-1068)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 670 ACGGGCGGTCGGATAAATCCAGTGTCCAGGGTGGGCGAGGATTCGCCATTCGGATC 729

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 DB 730 GGGCAGGCGATGGGATCCGGGCGGATCCGATCCGATCGGTTGGGGGTCAACCCCGTTTCAT 789

QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValArg 60  
 DB 790 ATCGGCGCTACCGCTTCCCTGGGTTGGTGTGTCACCAACGCGGCGGCGACGA 849

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 850 GTCCACCGGTCGGGAGCGCTCCGGCGGCAAGTCTCCGATCTCCACCGGCGAGTG 909

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 910 ATACCGCGGTGCGAGCGGCTCCGATCAACTCGGCGGCGGATCGGCGGCGGCGGCTTAAC 969

QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 970 GGGCATCATCCCGGTGACCTCATCTCGGTGACCTGGCAACCAAGTCCGGGCGGCGCGGT 1029

QY 121 ThrGlyAsnValThrLeuAlaGlu 128

Db 1030 ACGGGAACGTGACATTGGCGGAG 1053

RESULT 4  
 AX005788  
 LOCUS AX005788 1068 bp DNA linear PAT 24-AUG-2000  
 DEFINITION Sequence 907 from Patent WO909186.  
 ACCESSION AX005788  
 KEYWORDS AX005788.1 GI:9928795  
 SOURCE Mycobacterium tuberculosis  
 ORGANISM Mycobacterium tuberculosis  
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
 tuberculosis complex.  
 REFERENCE 1  
 AUTHORS Portnoi, D. and Guigueno, A.  
 TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors  
 comprising same and uses for diagnosing and preventing tuberculosis  
 Patent: WO 909186-A 907 25-FEB-1999;  
 JOURNAL PORTNOI DENIS (FR); GUIGUENO AGNES (FR)  
 FEATURES source  
 1..1068  
 /organism="Mycobacterium tuberculosis"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1773"  
 CDS  
 1..1068  
 /note="unnamed protein product; 50D"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC05178.1"  
 /db\_xref="GI:9928796"  
 /db\_xref="RENTREMBL:CAC05178"  
 /translations="MNSRRSLRWLLSVLAAVGLGLATAPQAAPPALSODRFAD  
 FPALPLDSAMVAQVGPQVYNNIKLGYNNAGAGTGVIDPNVVLNNHVIAGATD  
 INAFVSGGTGVYDVDTQDVAQLRGAGLPSAIGGVAVGSPVVMGNSG  
 GGGTTPRAVGRVVALGQTVQASDLSLTGARETLNGLIQDAAIQPDGSGPVVNGLGQ  
 VVGWNTAASDNFOLSGGGGFAPIGQAMAIAGQIESGGSPTHVIGPTAFILGIVD  
 NNGKARVQRVGSAPASGLSTGVTAVDGPINSATAMADALNGLHHPGDIIVT  
 WQTKSGGTRTGNVLAEGPPA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,428-42 Length: 1068  
 Score: 651.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AX005788 (1-1068)

QY 1 ThrAlaAlaSerAspAsnPhcGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
 DB 670 ACGGGCGGTCGGATAAATCCAGTGTCCAGGGTGGGCGAGGATTCGCCATTCGGATC 729

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlySerProThrValHis 40  
 DB 730 GGGCAGGCGATGGGATCCGGGCGGATCCGATCCGATCGGTTGGGGGTCAACCCCGTTTCAT 789

QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValArg 60  
 DB 790 ATCGGCGCTACCGCTTCCCTGGGTTGGTGTGTCACCAACGCGGCGGCGACGA 849

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 DB 850 GTCCACCGGTCGGGAGCGCTCCGGCGGCAAGTCTCCGATCTCCACCGGCGAGTG 909

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 DB 910 ATACCGCGGTGCGAGCGGCTCCGATCAACTCGGCGGCGGATCGGCGGCGGCGGCTTAAC 969

QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 DB 970 GGGCATCATCCCGGTGACCTCATCTCGGTGACCTGGCAACCAAGTCCGGGCGGCGCGGT 1029

QY 121 ThrGlyAsnValThrLeuAlaGlu 128

```

Db 970 GGGCATCATCCCGGTGACCTCATCTCGGTGACCTGGCAACCAAGTCGGGGGCGCACGCGT 1029
QY 121 ThrGlyAsnValThrLeuAlaGlu 128
Db 1030 ACAGGGAACGTGACATTCGCCGAG 1053

RESULT 5
BD274033 1143 bp DNA linear PAT 17-JUL-2003
LOCUS Sequences nucleic acid from polypeptides exportes from
DEFINITION mycobacteries, vector the component and applications at diagnostic
and the prevention from the tuberculose.
ACCESSION BD274033
VERSION 1
KEYWORDS 1 (bases 1 to 1143)
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1 (bases 1 to 1143)
AUTHORS Gicquel,B., Portnoi,D., Lim,E., Pellicic,V., Guigueno,A. and
Salmoniere,Y.G.D.L.
TITLE Sequences nucleic acid from polypeptides exportes from
mycobacteries, vector the component and applications at diagnostic
and the prevention from the tuberculose
JOURNAL Patent: JP 2002534956-A 257 22-OCT-2002;
INSTITUT PASTEUR
COMMENT OS Mycobacterium tuberculosis
PN JP 2002534956-A/257
PD 22-OCT-2002
PF 14-AUG-1998 JP 20000509849
PR 14-AUG-1997 FR 97/10404,11-SEP-1997 PR 97/11325 PI
BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
PC C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
C07K16/12,
PC C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
G01N33/53//
PC (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC SEQ
ID No. 50F
FH Key Location/Qualifiers
FT CDS (4)..(1140).

FEATURES
source
1..1143
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"

ORIGIN
Alignment Scores:
Pred. No.: 5,8e-42 Length: 1143
Score: 651.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x BD274033 (1-1143)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 745 ACGGCCGCGTCCGATAAATTCAGCTGTCCAGGTGGGGCGAGGATTCGCCATTCGATC 804

QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 805 GGGCAGCGCATGGCGATCCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 864

QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyValArg 60
Db 865 ATCGGGCCATCCCGCTTCCTCGGCTTGGGTGTGTGACAAACACCGCAACGGGCGACGA 924

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

```

```

Db 925 GTCCAAACGCGTGGTCCGGAGCGCTCCGGCGCGCAAGTCTCGGCATCTCCACCGCGCAGTG 984
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 985 ATCACCAGCGGTCCGAGCGGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCTTAAC 1044
QY 101 GlyHisHisProGlyAspValIleSerValThrTyrGlnThrIleThySerGlyGlyThrArg 120
Db 1045 GGGCATCATCCCGGTGACGTCATCTCGTGACCTGGCAACCAAGTCGGGGCGCACGCGT 1104
QY 121 ThrGlyAsnValThrLeuAlaGlu 128
Db 1105 ACAGGGAACGTGACATTCGCCGAG 1128

RESULT 6
AX005790 1143 bp DNA linear PAT 24-AUG-2000
LOCUS Sequence 909 from Patent WO9909186.
DEFINITION AX005790
ACCESSION AX005790
VERSION AX005790.1 GI:9928797
KEYWORDS Mycobacterium tuberculosis
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1
AUTHORS Portnoi,D. and Guigueno,A.
TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors
comprising same and uses for diagnosing and preventing tuberculosis
JOURNAL Patent: WO 9909186-A 909 25-FEB-1999;
PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES
source
1..1143
/organism="Mycobacterium tuberculosis"
/mol_type="unassigned DNA"
/db_xref="taxon:1773"
<4..1143
/notes="unnamed protein product; SEQ ID NO 50F"
/codon_start=1
/transl_table=11
/protein_id="CAC05179.1"
/db_xref="GI:9928798"
/db_xref="RENTREMBL:KAC05179"
/translacion="KNPAARTLKAGTISDSDYPTGGYGNMSRRSLRWSLLSLVLA
VGLGLATAQAAPPAPQRFADFPALPDPDSAMVAQVPOVYNTVGLYNNAYGA
GTGIVIDPENGVLNTNHHVIAGATDINAFSGSGQTYGVGVGVYDRTQDVAVLQRGAG
GLPSAAICGGVAVGEPVVMGNSGGCGGTFRVGRVVALGQTVQASDSLTCAEFTLN
GLIQFDAIQGDSGGPVVNGLGQVGMNTAASDNFQLSGGGQGFPAIPGQMAIAGQ
IRSGGSPTHVIGTFAFLGIVGVDDNNGARVORVVGVSAPASLSLSTGDIVITAVDGA
PINSATAMADALNHGHPGDIVSVTWTKSGGTETGNVTLAEGPPA"

Alignment Scores:
Pred. No.: 5,8e-42 Length: 1143
Score: 651.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AX005790 (1-1143)

QY 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 745 ACGGCCGCGTCCGATAAATTCAGCTGTCCAGGTGGGGCGAGGATTCGCCATTCGATC 804

QY 21 GlyGlnAlaMetAlaIleAlaGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 805 GGGCAGCGCATGGCGATCCGGGGCCAGATCCGATCGGTGGGGGTACCCACCGTTTCAT 864

QY 41 IleGlyProThrAlaPheLeuGlyValValAlaAspAsnGlyValArg 60

```

```

Db      865 ATCGGGCTACCGCTTCTCGGCTTGGGTGTGTGGACAAACGCGACGCGACGA 924
Qy      61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db      925 GTCCACGCGGTGTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGCACGCG 984
Qy      81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db      985 ATCACCAGCGTTCGACGCGCTCCGATCACTCGGCCACCGGATGCGGACGCGCTTAAC 1044
Qy      101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlySerGlyGlyThrArg 120
Db      1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGACGCGT 1104
Qy      121 ThrGlyAsnValThrLeuAlaGlu 128
Db      1105 ACAGGGAACGTGACATTGGCCGAG 1128

RESULT 7
BD251322
LOCUS      BD251322
DEFINITION Fused protein of Mycobacterium tuberculosis antigen and utilization thereof.
ACCESSION BD251322
VERSION    BD251322.1 GI:33061092
KEYWORDS   JP 2002510494-A/1.
SOURCE     synthetic construct
ORGANISM   artificial sequences.
REFERENCE  1 (Bases 1 to 2287)
AUTHORS    Skeiky, Y.A.W., Alderson, M. and Neto, A.C.
TITLE      Fused protein of Mycobacterium tuberculosis antigen and utilization thereof
JOURNAL    Patent: JP 2002510494-A 1 09-APR-2002;
COMMENT    CORIXA CORP
OS         Artificial Sequence
PN         JP 2002510494-A/1
PD         09-APR-2002
PF         07-APR-1999 JP 2000542460
PR         07-APR-1998 US 09/056556,30-DEC-1998 US 09/223040 PI
YASIR A W SKEIKY, MARK ALDERSON, ANTONIO CAMPOS NETO PC
CI2N15/09, A61K39/04, A61K48/00, A61P31/04, C07K14/35, C07K19/00, PC
CI2P21/02
PC         CL2N15/00
CC         Description of Artificial Sequence: tri-fusion protein Ra12- CC
           TBH9-Ra3s
           (designated Mtb32A)
CC         n = g, a, c or t
CC         n = g, a, c or t
CC         n = g, a, c or t
PH         Key Location/Qualifiers
PT         modified base (30)
FT         modified_base (33)
FT         modified_base (42)... (2231)
FT         modified_base (2270).
FT         Location/Qualifiers
           1..2287
           /organism="synthetic construct"
           /mol_type="genomic DNA"
           /db_xref="taxon:32630"

ORIGIN
Alignment Scores:
Pred. No.: 1.16e-41 Length: 2287
Score: 651.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x BD251322 (1-2287)

```

```

Qy      1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db      63 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 122
Qy      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db      123 GGGCAGCGGATGCGGATCGCGGCCAGATCCGATCCGGTGGGGGTGTCACCCACCGTTTCA 182
Qy      41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db      183 ATGGGCTTACCCCTTCTCGGCTTGGGTGTGTTCGACAAACGCGAACCGCGCACGA 242
Qy      61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
Db      243 GTCCACGCGGTGTCGGAGCGCTCCGCGCGCAAGTCTCGGCATCTCCACCGCGCACGCG 302
Qy      81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db      303 ATCACCAGCGTTCGACGCGCTCCGATCACTCGGCCACCGCGATGGCGGACGCGTTAAC 362
Qy      101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlySerGlyGlyThrArg 120
Db      363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAACCAAGTCGGCGCGACGCGT 422
Qy      121 ThrGlyAsnValThrLeuAlaGlu 128
Db      423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 8
AR303127
LOCUS      AR303127
DEFINITION Sequence 1 from patent US 6544522.
ACCESSION AR303127
VERSION    AR303127.1 GI:31691855
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (Bases 1 to 2287)
AUTHORS    Skeiky, Y., Alderson, M. and Campos-Neto, A.
TITLE      Fusion proteins of mycobacterium tuberculosis antigens and their uses
JOURNAL    Patent: US 6544522-A 1 08-APR-2003;
FEATURES   Location/Qualifiers
           source
           1..2287
           /organism="unknown"
           /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.16e-41 Length: 2287
Score: 651.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR303127 (1-2287)
Qy      1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
Db      63 ACGGCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCGATC 122
Qy      21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db      123 GGGCAGCGGATGCGGATCGCGGCCAGATCCGATCCGGTGGGGGTGTCACCCACCGTTTCA 182
Qy      41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db      183 ATGGGCTTACCCCTTCTCGGCTTGGGTGTGTTCGACAAACGCGAACCGCGCACGA 242
Qy      61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

```

```
Db 243 GTCACCGCGTGGTCCGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGGACGTG 302
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCGCGTCCAGCGGCGTCCGATCAACTACCGCCACCGGATGGCGGACGCGCTTAAC 362
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 363 GGGCATCATCCCGGTGGAGCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 422
Qy 121 ThrGlyAsnValThrLeuAlaGlu 128
Db 423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 9
LOCUS AR403735 2287 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6627198.
ACCESSION AR403735
VERSION AR403735.1 GI:40151411
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Unclassified.
TITLE 1 (bases 1 to 2287)
JOURNAL Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
Fusion proteins of Mycobacterium tuberculosis antigens and their
uses
JOURNAL Patent: US 6627198-A 1 30-SEP-2003;
FEATURES
Location/Qualifiers
source
1..2287
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.16e-41 Length: 2287
Pred. No.: 651.00 Matches: 128
Score: 851.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR403735 (1-2287)
Qy 1 ThrAlaIleSerAspAsnGlyLeuSerGlnGlyGlnGlyPheAlaIleProIle 20
Db 63 ACAGCGCGTCCGATCACTTCCAGCTGTCCAGGCGTGGCGAGGATTCGCCATTCGGATC 122
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db 123 GGGCAGCGGATGGCGATCCGCGGCCAGATCCGATCGGTGGGGGTCACCCACCGTTCAT 182
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
Db 183 ATCGGCGCTACCGCTTCTCGGTGGTGTTCGACAAACACGCAACGCGCACGCA 242
Qy 61 ValGlnArgValValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80
Db 243 GTCCAAACGCGGTGGAGCGCTCCGCGGCGCAAGTCTCGCATCTCCACCGCGGACGTG 302
Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
Db 303 ATCACCGCGTCCAGCGGCGCTCCGATCACTCGGCCACCGCATGGCGGACGCGCTTAAC 362
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
Db 363 GGGCATCATCCCGGTGGAGCTATCTCGGTGACCTGGCAACCAAGTCGGGCGGACGCGT 422
Qy 121 ThrGlyAsnValThrLeuAlaGlu 128
Db 423 ACAGGGAACGTGACATTGGCCGAG 446
```

```
RESULT 10
LOCUS AE006925 14029 bp DNA linear BCT 27-APR-2001
DEFINITION Mycobacterium tuberculosis CDC1551, section 11 of 280 of the
complete genome.
ACCESSION AE006925 AE000516
VERSION AE006925.1 GI:13879610
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis CDC1551
Mycobacterium tuberculosis CDC1551
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1 (bases 1 to 14029)
Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 14029)
Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
source
1..14029
/organism="Mycobacterium tuberculosis CDC1551"
/mol_type="genomic DNA"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
complement(99..533)
/gene="WT0129"
complement(99..533)
/gene="WT0129"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44353.1"
/db_xref="GI:13879611"
/translation="MGSEFDKLRFAQSPVARLATSTPDGPHLVVPVVFALCARRPAEA
TGADVITAVDAKKTKTQRLRLRLANLHNPRASVLVDSYADDWTQLWVRADGVAAIH
RDGVMRAAYKLRKATQYQSVPLNGFVIAIAVQRWASRA"
complement(643..1062)
/gene="WT0130"
complement(643..1062)
/gene="WT0130"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44354.1"
/db_xref="GI:13879612"
/translation="MASSSSPLLRHCVSDDQVTVVGFDDGLGKTARRIAALVYQRA
IFLNDRTVAHSGDDAVLGHAVLPVGPDPDPDLPHASSMYSLGMCQSVNGRPFDAIA
LVSVRLCHVQDPTDSCGRDRPGQLPCAPLDYHRH"
1047..1415
/gene="WT0131"
1047..1415
/gene="WT0131"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="DNA-binding protein, CopG family"
/protein_id="AAK44355.1"
```



```
/db_xref="GI:13879613"
/translation="MTKKPKNADYVIGDDVEVDVLDKOEYVVDGERLTDREVSQM
ASESLRLAKSRKREANLIFGGKSLGSGASHSPAVQVWSKATHAKLKLARSRKMSVSKL
LRPVLDFBQVQTRGILPRR"
1682..3367
/gene="WT0132"
CDS
1682..3367
/gene="WT0132"
/translation="PE PGRS family protein"
/notes="similar to GB:AL23456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="PE PGRS family protein"
/protein_id="AAK44356.1"
/db_xref="GI:13879614"
/translation="MEAGGAGGRCVSVFVAPRIVVAAATDLAGISASAAANA
AAAATTAVLAGADENVSRAIALFSGHAQAYALSQAAPHQFQVTLGAGGAGA
AAEQVEQLLAALIGTALLRPLIGNADGAPGTQAGAGGILYNGGNGGSGA
AQAGAGGAGLIGHGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CHGAGGAGLWAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
QAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
LIGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CHLPNGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
PGNGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
3519..4586
/gene="WT0133"
CDS
3519..4586
/gene="WT0133"
/notes="identified by match to protein family HMW"
/codon_start=1
/transl_table=11
/product="serine protease, putative"
/protein_id="AAK44357.1"
/db_xref="GI:13879615"
/translation="MSNRRRLRWSLLSLVLAAGLGLATAPAAQAPALSDQRFAD
FALPLDPSAMVAVQVPOVNTKLGYNVAGTGVIDPVGVLVNNHVIAGATD
INAFVSGQTTGVNDVQDRTQDVANLQRCAGLPSAAGGVAGVGPVVMAGNSG
GQGTTPVAGVVDVQVQASDLTGAEFTLGLIIPDAIAPQDGGGVPVWGLCQ
VVMNTAASDNTQLSGGGGFAIPQAMAIAGQSGGSGPTVHIGTAPLGLGVVD
NNGNGARVQVWVSGAPASLGISTGVITAVDGPATNSATAMADALNGHPGDVIST
WOTKSGCTRTGNTVLAEGPPA"
4695..6500
/gene="WT0134"
CDS
4695..6500
/gene="WT0134"
/notes="similar to GP:2808807; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="alpha-amylase family protein"
/protein_id="AAK44358.1"
/db_xref="GI:13879616"
/translation="MNEAHSVEHPVQSGHVEGVEHPDADKDFGSAALPADTWP
KXAFVFLVAFADSDGDLGLIDLQWGLDYLQWGLDYLQWGLDYLQWGLDYLQW
IRDFYKVLPEFGVDDFVALVDAHRRGIRITDLMNHTSHPFQSRDPPGYP
GDYVNSDTSERTDARIIPVDESNWSFDPVRQFYWHRRFHSHPDLYNDPAVQE
AMTDVRFWGLGIDPGLDAVPLPERSGNCENPETHAFKVRKVVDDBFQGV
LLEAQMGGVVEYEDNTGDECHMAFHPMPRIEMVRRSRFPISLIIQTP
PLDQMQGIFLPHHDELTLEMTDEEDIMABYAKDPRKRVIRRLAPLIND
RQTEIFALLSLPGSVLYGDEIGMVDVWLDGDRGVRIPQMTPDRNAGFSTAN
PGLIYLPSPQDPVYQAVNVEARDTSTSLNFTRLMVAHRRHPAFAGFQELGG
SNPSVLAYVQVAGDDGDTVLCVNLRSFPQPIELDQWNTYTPVELTGHVEFRIG
QVPVLLPLCGHGFWEQLTTHEVGAPTCGGERRL"
6603..7970
/gene="WT0135"
CDS
6603..7970
/gene="WT0135"
/notes="similar to GP:2808801; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="pep2 protein"
```

```
/protein_id="AAK44359.1"
/db_xref="GI:13879617"
/translation="WTRSDTLATKLPSWDLPRQWYAGNRRELATVPGVVVALRHN
LDLVLVDVYTDGATERYQVLVWGDVEPASEYTKAAICVADDRGTGFDALVDVAGPQF
LLSLVSSACCTSTGEVTFEPEDVELPRAQPRVCDAEQSNSTSVIPDRRAILKVPFR
RVSSGNDIENRVITRAGNPHVARLLGACVQFGRPNRSPDATALAVAGMTVEYANA
AEGMATASVSVLFAEGDLYAHEVPTIEQFQKLAABEITVQRVHGDLHGLGVKRTIPES
PYDRMLARLSTVAVVPELREYATTEQFQKLAABEITVQRVHGDLHGLGVKRTIPES
WLLIDFEGEPQGLDERAPDRLDVGALVRSFYAAYGPLVDQATDKQLAARREW
VERNRAAFCDGVAVASGIDPRDSALLGAVELDKAVYETGYETRHPRPGWLPIPLRSIA
RLTAS"
8038..8817
/gene="WT0136"
CDS
8038..8817
/gene="WT0136"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44360.1"
/db_xref="GI:13879618"
/translation="MOREYDGEARLSWLAALAGILGATATFTHSAGYFVTWMTGNSQ
RAVLGLFGDDAMSVTASLILFFVAGVVIASVCRHFHWAHPHGTVLITSLIPAA
RVDMLGGWHESMLDFVPLIFVVGIGALNTSFVKDGEVSLVYVYVTLVKNQOIE
RHLAGKVDWLGVFYLHASFVLAAGAGSISMTVTGPQMLAVAAVVCATTTGYTLH
ADRRGLVQKRPQPKRLPRALRRGLDSGTSTPTAANYGSS"
complement(8949..9989)
/gene="WT0137"
CDS
complement(8949..9989)
/gene="WT0137"
/notes="similar to SP:P31953 GB:X57229 PID:48828
PID:156734 PID:1877254; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="esterase, putative, antigen 85-C"
/protein_id="AAK44361.1"
/db_xref="GI:13879619"
/translation="MGLVAMTFFEOVRRLRSAAITLPRRLAIAAMGAVLYGLVGTFF
GGTACAGFRRPGLPVEYLOVPSASMGRIKVOFGGPHAVYLLDGLRAQDDYNGWD
INTPAFEYQSGLSVIMPVGSSSEYTDVQPSQNGONYTKWETFLREMPAMLOQ
ANKGSYTNAAAGLSMSGSSALILAAIYQFQPYAASLSGFLNPSGHWFTLGLAM
NDGSGVSNAMWGFSPFAWRNDPMQIPRLVANNTRIWYCGNGTFFSLDGNIPFA
KFLGLTLRTNQTFRDTYAADGGNGVFNFPNGTHSWPYWNBQLVAMKADIQHLVNG
ATPPAAPAAPAA"
10218..10673
/gene="WT0138"
CDS
10218..10673
/gene="WT0138"
/notes="similar to SP:P08634 PID:46222; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="modulation protein N-related protein"
/protein_id="AAK44362.1"
/db_xref="GI:13879620"
/translation="ThraAaAaSerAapAapPheGlnLeuSerGlnGlyGlnGlyPheAaAaIleProfile 20
4189 ACGGCGCGTCCGATAACTTCAGCTGTCACAGGTGGCAGGATTCGCCATCCGATC 4247
21 GlyGlnAaMetAaIaAaAaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
```

Alignment Scores:

Align. No.:	7,2e-41	Length:	14029
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1	Indels:	0
DB:		Gaps:	0

US-09-684-215B-23 (1-128) x AE006925 (1-14029)



Db 4248 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTTCAT 4307  
QY 41 lleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyValAlaArg 60  
Db 4308 ATCGGCGCTACCGCTTCTCGGCTTGGTGTTCGACACACGCGAACGGCGCAGCA 4367  
QY 61 ValGlnArgValGlySerAlaProAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 4368 GTCCAAACGCGTGGTGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGACGTG 4427  
QY 81 lleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 4428 ATCACCAGCGGTGACGGCGCTCCGATCACTCGGCACCGCGATCGGACGCGCTTAAC 4487  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyClyThrArg 120  
Db 4488 GGGCATCATCCCGGTGACGTCTATCTCGGTGACCTGGCAACCAAGTCGGCGGCGACGGT 4547  
QY 121 ThrGlyAsnValThrIleuAlaGlu 128  
Db 4548 ACAGGAACGTGACATTGGCCGAG 4571  
RESULT 11  
BX842572  
ID BX842572 standard; circular genomic DNA; PRO; 341957 BP.  
XX  
AC BX842572; AL021427; AL021428; AL021926; AL021927; AL021928; AL021929;  
AC AL021930; AL123456; Z74410; Z80233; Z80775; Z86089; Z92665; Z92770; Z96071;  
AC Z97050;  
XX  
SV BX842572.1  
XX  
DT 21-NOV-2003 (Rel. 77, Created)  
DT 21-NOV-2003 (Rel. 77, Last updated, Version 1)  
XX  
DE Mycobacterium tuberculosis H37Rv complete genome; segment 1/13  
XX  
KW complete genome.  
XX  
OS Mycobacterium tuberculosis H37Rv  
XX  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;  
OC Mycobacterium tuberculosis complex; Mycobacterium tuberculosis.  
XX  
XX  
[1]  
RX MEDLINE; 98295987.  
RX PUBMED; 9634230.  
RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry III C.E., Teklaia F., Badcock K.,  
RA Bigham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K.,  
RA Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jorgensen K.,  
RA Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J.,  
RA Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S.,  
RA Squares S., Squares R., Sultston J.E., Taylor K., Whitehead S., Barrall B.G.,  
RA "Deciphering the biology of Mycobacterium tuberculosis from the complete  
RT genome sequence";  
RX Nature 393:537-544 (1998).  
XX  
[2]  
RX PUBMED; 12368430.  
RX Camus J.C., Pryor M.J., Medigue C., Cole S.T.;  
RA "Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv";  
RL Microbiology 148:2967-2973 (2002).  
XX  
[3]  
RP 1-341957  
RA Parkhill J.;  
RT Submitted (11-JUN-1998) to the EMBL/GenBank/DBJ databases.  
RL Submitted on behalf of the Mycobacterium tuberculosis sequencing and  
RL mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton,  
RL Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut  
RL Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail:

parkhill@sanger.ac.uk  
Notes: Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.  
(URL, [http://www.sanger.ac.uk/Projects/M\\_tuberculosis/](http://www.sanger.ac.uk/Projects/M_tuberculosis/))  
Key Location/Qualifiers  
source 1. 341957  
/db\_xref="taxon:83332"  
/mol\_type="genomic DNA"  
/organism="Mycobacterium tuberculosis H37Rv"  
/strain="H37Rv"  
1. 1524  
/evidence=EXPERIMENTAL  
/notes="RV0001, (MT0001, MT0029.01, P49993), len: 507 aa.  
dnaA, chromosomal replication initiator protein (see citations below), equivalent to other Mycobacterial CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g. P46388 [DNAA\_MYCLE from Mycobacterium leprae (502 aa); Q9L7L7 [DNAA\_MYCPA from Mycobacterium paratuberculosis (509 aa); P49990 [DNAA\_MYCAV from Mycobacterium avium (508 aa); P49992 [DNAA\_MYCSM from Mycobacterium smegmatis (504 aa); etc. Also highly similar to others except in N-terminus e.g. Q9ZHT5 [DNAA\_STRCH CHROMOSOMAL REPLICATION INITIATOR PROTEIN from Streptomyces chrysomallus (624 aa); Q9ZH76 [DNAA\_STRRE from Streptomyces reticuli (643 aa); DNAA\_ECOLI [P03004 [B3702 chromosomal replication initiator protein from Escherichia coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0, (43.2% identity in 389 aa overlap); etc. Contains P80017 ATP/GTP-binding site motif A (P-loop) and P80108 DnaA protein signature. BELONGS TO THE DNAA FAMILY. Note that the first base of this gene has been taken as base 1 of the Mycobacterium tuberculosis H37Rv genomic sequence."  
/transl\_table=11  
/function="PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION; IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS (DNAA BOX): 5'-TTATC(C/A)A(C/A)A-3'. DNAA BINDS TO ATP AND TO ACIDIC PHOSPHOLIPIDS. DNAA PROTEIN BINDS THE ORIGIN OF REPLICATION (oriC), ATP AND ADP, AND EXHIBITED WEAK ATPase ACTIVITY."  
/gene="dnaA"  
/locus\_tag="RV0001"  
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"  
/protein\_id="CAA16238.1"  
/translation="MTDDPGSGFTTVNNAVVSSELNGLPKVDGSSDANLSAPLTPOOR AWWLNLQPLTIVEGFALLISVFSSVQNEIEHRLAPITDALSRELGHQIQGVRIAPPA TDEADDTVPSENPATTSPTTDDNDEIDSAARGNDQNSPYSYTERPHNTDSATA GVTSLNREYTFDTFVIGASRFHAAALIAEAPARAYNPLFIWGESGLGTHLHAAG NYAQLPFGMKVKVSTEEFTNDINSRDDKRVAKRSYRDVLLVDDIQFTEGKEG IQEEFFHTLHNANKQIVISSRPPKQLATLDRLETRFEMGLITDQPPLETRIA ILRKAQMERLAVPDVDELIIASSIRNEIREGALIRVAFASINKTPIDKALAEIVL RDLADANTMQISAATIAATAFYDFTVEELRGKTRALAQKRIARQIAYLCRELTDL LPKIQAFGRDHTTVMYAQRKILSEMAERREVDHVKELTRIRQKSR"  
622..645  
/note="P800017 ATP/GTP-binding site motif A (P-loop)"  
misc\_feature 1384..1440  
/note="P801008 DnaA protein signature"  
misc\_feature 2052..3260  
/note="RV0002, (MTV029.02, MTCV10H4.0), len: 402 aa. dnaN, DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations below), equivalent to other Mycobacterial DNA POLYMERASES III BETA CHAIN e.g. NP 301130.1 [NC 002677 from Mycobacterium leprae (399 aa); Q9L7L6 [DP38 MYCPA from Mycobacterium avium subsp. paratuberculosis (399 aa); P2851 [DP38 MYCSM from Mycobacterium smegmatis (397 aa); etc. Also highly similar to others e.g. P27903 [DP38 STRCO DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity

FT in 337 aa overlap; P21174|DP3B\_MICU from Micrococcus  
 FT luteus (310 aa); P52023|DP3B\_SYN7 from Synechococcus sp.  
 FT strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in  
 FT neighbouring cosmid MTCV10H4.01."  
 FT /transl table=1  
 FT /EC number="2.7.7.7"  
 FT /function="DNA POLYMERASE III IS A COMPLEX, MULTICHAIN  
 FT ENZYME RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN  
 FT BACTERIA. THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5'  
 FT EXONUCLEASE ACTIVITY. THE BETA CHAIN IS REQUIRED FOR  
 FT INITIATION OF REPLICATION ONCE IT IS CLAMPED ONTO DNA, IT  
 FT SLIDES FREELY (BIDIRECTIONAL AND ATP-INDEPENDENT) ALONG  
 FT DUPLEX DNA [CATALYTIC ACTIVITY: N deoxynucleoside  
 FT triphosphate + N diphosphate + [DNA]n"]  
 FT /gene="dnaN"  
 FT /locus tag="rv0002"  
 FT /product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA  
 FT NUCLEOTIDYLTRANSFERASE)"  
 FT /protein\_id="CA116239.1"  
 FT /translation="MDAATRVGLTDLTFLRLRESFADAVSWAKNLPARPAPVPLSGV  
 FT LRTGSDNLATISGDFEVSAAQVGAIEIVSGVLVSGRLSDITRALPNKVDVHVHG  
 FT NRVALTGNAARFSLPTMPVEDYPTLPLEETGLLPAELFAIAISQVATAAGRDITLPM  
 FT LTGRVRLGETVVLATDRFLAVRELKWSASPDIEAAVLVPKTLAERAKAGIGGS  
 FT DVRLSITGFGVKGDLGISGNGKSTRLLDARFPKRLQPLTFHTAVATMDVABLI  
 FT EAKLVADRGAVQRMFPADGVSRLAGADDVGRASDLVDVYAGEPLTIAFNPYLI  
 FT TDGLSUSERSVSFGFTTAGKPAALLRPVSGDDRPVAGLNGNGPFPFAVSTDYTYLLMPVR  
 FT LPP"  
 FT 3280..4437  
 FT /evidence=EXPERIMENTAL  
 FT /note="rv0003. (MTCV10H4.01), len: 385 aa. recF, DNA  
 FT replication and repair protein (see citations below), and  
 FT equivalent to others Mycobacterial DNA replication and  
 FT repair proteins e.g. NF\_301131.1|NC\_002677 from  
 FT Mycobacterium leprae (385 aa); Q9L7L5|REFC MYCPA from  
 FT Mycobacterium avium subsp. paratuberculosis (385 aa);  
 FT P50916|REFC MYCSM from Mycobacterium smegmatis (384 aa);  
 FT etc. Also highly similar to others e.g. P36176|REFC STRCO  
 FT DNA REPLICATION AND REPAIR PROTEIN from Streptomyces  
 FT coelicolor (373 aa); NP\_440892.1|NC\_000911 from  
 FT Synechocystis sp. strain PCC 6803 (384 aa);  
 FT NP\_469352.1|NC\_003212 from Listeria innocua (370 aa); etc.  
 FT Contains PS00017 ATP/GTP-binding site motif A (P-loop),  
 FT PS00617 RecF protein signature 1, and PS00618 RecF protein  
 FT signature 2. BELONGS TO THE RECF FAMILY."  
 FT /transl table=1  
 FT /function="THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM  
 FT AND RECOMBINATION: IT IS REQUIRED FOR DNA REPLICATION AND  
 FT NORMAL SOS INDUCIBILITY. RECF BINDS PREFERENTIALLY TO  
 FT SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS TO BIND ATP."  
 FT /gene="recf"  
 FT /locus tag="rv0003"  
 FT /product="DNA REPLICATION AND REPAIR PROTEIN RECF  
 FT (SINGLE-STRAND DNA BINDING PROTEIN)"  
 FT /protein\_id="CA02424.1"  
 FT /translation="MVYRHGLRDFRSWACVDELEHPGRTVFVGNPGYKTLIALAW  
 FT STTGSHRVSADLPLISGCTDRAVISTVYVNDGRCANDELTATGHWKALNRSVRS  
 FT TRDVGVRLAVFAPDLGVGDPAARRRYLDLAIIVRPAIAVRAIYERVLKQRTA  
 FT LKSVFGRVNGDGVFDLTLEWDSHGAELAAIDLVNQLAPGVKKYQQLLAP  
 FT SRSAISGIRASVDYSEQSDIDQLAAALAAARDAELRGVGLVGHPRDDL  
 FT LRLGQPAKGFASGEASLVALRLAQLLRVDCGPEVLLDDVFAELDWYMRRLA  
 FT TAASASQVLTVTALELDIPAGWDARVHIDVRADDTGSMVLP"  
 FT 3367..3390  
 FT /note="PS00017 ATP/GTP-binding site motif A"  
 FT 3634..3690  
 FT /note="PS00617 Recf protein signature 1"  
 FT 4243..4296  
 FT /note="PS00618 Recf protein signature 2"  
 FT 4434..4497  
 FT /evidence=EXPERIMENTAL  
 FT /note="rv0004. (MTCV10H4.02), len: 187 aa. Conserved  
 FT hypothetical protein (see Salazar et al., 1996), highly  
 FT similar, but longer 21 aa in N-terminus, to  
 FT AAF33696.1|AF222789 unknown protein from Mycobacterium

FT avium subsp. paratuberculosis (166 aa); and highly similar  
 FT to NP\_301132.1|NC\_002677 conserved hypothetical protein  
 FT from Mycobacterium leprae (189 aa); S70990 hypothetical  
 FT protein from Mycobacterium smegmatis (194 aa). Also highly  
 FT similar, except in N-terminal part, to  
 FT  
 FT Alignment Scores: Length: 341957  
 FT Pred. No.: 1.78e-39 Matches: 128  
 FT Score: 651.00  
 FT Percent Similarity: 100.00% Conservative: 0  
 FT Best Local Similarity: 100.00% Mismatches: 0  
 FT Query Match: 100.00% Indels: 0  
 FT DB: 15 Gaps: 0  
 FT  
 FT US-09-684-215B-23 (1-128) x BX842572 (1-341957)  
 FT  
 FT QY 1 ThrAlaAlaSerAAsenPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
 FT DB 151817 ACGCCCGCTCCGATTAATCTCCAGCTGTCCAGGGTGGCCAGGATTCGCCATTCGCATC 151876  
 FT QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleAArgSerGlyGlyGlySerProThrValHis 40  
 FT DB 151877 GGGCAGCGGATGGGATCGGGGCCAGATCCGATCGGTGGGGGTCCACCGCGTTTCAAT 151936  
 FT QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
 FT DB 151937 ATCGGGCTTACCCCTTCCTCGCTTGGGTGTGTGACAAACACGCGCAACGCGCACGA 151996  
 FT QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 FT DB 151997 GTCCACGCGTGGTTCGGGAGCGTCCGGGGCAGATCTCGGCATCTCCACCGCGCGCTG 152056  
 FT QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 FT DB 152057 ATCACCAGCGTGCACGGCGTCCGATCAACTCGGCACCGCATCGCGACGCGCTTAAC 152116  
 FT QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 FT DB 152117 GGGCATCATCCGGTGACGTCATCTCGTGACCTGCGAACACCACTCGGCGCGCGCT 152176  
 FT QY 121 ThrGlyAsnValThrIleAlaGlu 128  
 FT DB 152177 ACAGGGAACGTGACATTTGGCCGAG 152200  
 FT  
 FT RESULT 12  
 FT BX248334 343050 bp DNA linear BCT 11-JUN-2003  
 FT BX248334  
 FT LOCUS Mycobacterium bovis subsp. Bovis AF2122/97 complete genome; segment  
 FT DEFINITION 1/14.  
 FT ACCESSION BX248334 BX248333  
 FT VERSION BX248334.1 GI:31616762  
 FT KEYWORDS complete genome.  
 FT SOURCE Mycobacterium bovis subsp. bovis AF2122/97  
 FT ORGANISM Mycobacterium bovis subsp. bovis AF2122/97  
 FT Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 FT Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
 FT tuberculosis complex.  
 FT  
 FT REFERENCE  
 FT AUTHORS  
 FT Garnier, T., Eiglmeyer, K., Camus, J.-C., Medina, N., Mansoor, H.,  
 FT Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsemp, C., Simon, S.,  
 FT Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L.,  
 FT Wheeler, P., Parkhill, J., Barrell, B.G., Cole, S.T., Gordon, S.V. and  
 FT Hewinson, G.  
 FT The complete genome sequence of Mycobacterium bovis  
 FT Online Publication  
 FT FNAS 10.1073/pnas.1130426100 ( Microbiology )  
 FT REFERENCE 2 (bases 1 to 343050)  
 FT Garnier, T.  
 FT Direct Submission  
 FT Submitted (24-WAR-2003) Garnier T., Unite de Genetique Moleculaire  
 FT Bacterienne Institut Pasteur 28 rue du Dr Roux 75724 PARIS cedex  
 FT 15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the  
 FT Mycobacterium bovis sequencing teams, TB Research Group, Veterinary

Laboratories Agency Weybridge, Woodham Lane, New Haw, Addlestone, Surrey, K11 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. P74 Annotation, Genopole, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France. Unité de Génétique Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

# FEATURES

Location/Qualifiers

1. 343050  
 /organism="Mycobacterium bovis subsp. bovis AF2122/97"  
 /mol\_type="genomic DNA"  
 /strain="AF2122/97"  
 /db\_xref="taxon:233413"  
 1. 1524  
 /gene="dnaA"  
 /locus\_tag="Mb0001"  
 1. 1524  
 /gene="dnaA"  
 /locus\_tag="Mb0001"  
 /note="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001.  
 len: 507 aa, from Mycobacterium tuberculosis strain H37Rv, (99.6% identity in 507 aa overlap). dnaA, chromosomal replication initiator protein (see citations below), equivalent to other Mycobacterial CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g. P46388|DNAA MYCLE from Mycobacterium leprae (502 aa); Q9L7L7|DNAA MYCPA from Mycobacterium paratuberculosis (509 aa); P49990|DNAA MYCAV from Mycobacterium avium (508 aa); P49992|DNAA MYCSM from Mycobacterium smegmatis (504 aa); etc. Also highly similar to others except in N-terminus e.g. Q92H75|DNAA\_STRCH CHROMOSOMAL REPLICATION INITIATOR PROTEIN from Streptomyces chrysomallus (624 aa); Q92H76|DNAA\_STRRE from Streptomyces reticuli (643 aa); DNAA ECOLI|P03004|B3702 chromosomal replication initiator protein from Escherichia coli strain K12 (467 aa), FASTA scores: Opt: 986, E(): 0, (43.2% identity in 389 aa overlap); etc. Contains P500017 ATP/GTP-binding site motif A (p-loop) and P501008 DnaA protein signature. BELONGS TO THE DNAA FAMILY. Note that the first base of this gene has been taken as base 1 of the Mycobacterium bovis genomic sequence."  
 /codon\_start=1  
 /evidence=experimental  
 /transl\_table=1  
 /product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"  
 /protein\_id="CAD92863.1"  
 /db\_xref="GI:31616763"  
 /translation="MTDDPGSGFTTWNVAVSELDGPKVDKDPSSDANLSAPLTPQQ  
 RAWLNVOPLTIVEGFALLVPSPFVONEIERHLRAPITDALSRLGHQIQGVRIAP  
 PATDEADDTVPSPENPATTSPDTTNDIDSDAARGDNQSPSYTERPNTDS  
 ATAGVTSLNRYTPTTVIGASNRFAHAALAAEAPAYNPLFTWGSGLGKTHLL  
 HAAGNYAQLFPGMRVYVSTEEFTNDINSLRDRKVAFKRSYRDVLLVDDQFI  
 EGKGIQEPPTHTLNANKQIVISSDRPPQLATLEDRLETRFWGLITDVDPPE  
 LETRIALRKAQWERLAIPODVLLELIASSIERIRELGALIRVTFAPASLNTPIDK  
 ALAEIVLDLADANTWQISNACTMTATASYFTTVEELRGPKTEALAQSQIAMYL  
 CRELTDLSPKIGQAFGRDHTTWNVAQRKILSEARREVFHVKELTTRIRQSRK"  
 2052..3260  
 /gene="dnaN"  
 /locus\_tag="Mb0002"  
 2052..3260  
 /gene="dnaN"  
 /locus\_tag="Mb0002"  
 /EC\_number="2.7.7.7"  
 /note="Mb0002, dnaN, len: 402 aa. Equivalent to Rv0002.  
 len: 402 aa, from Mycobacterium tuberculosis strain H37Rv, (99.8% identity in 402 aa overlap). dnaN, DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations below), equivalent to other Mycobacterial DNA POLYMERASES III BETA CHAIN e.g. NP 301130.1|NC 002677 from Mycobacterium leprae (399 aa); Q9L7L6|DP3B MYCPA from Mycobacterium avium subsp. paratuberculosis (399 aa); P52851|DP3B MYCSM from Mycobacterium smegmatis (397 aa); etc. Also highly similar to others e.g. P27503|DP3B\_STRCO DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor (376 aa), FASTA scores: Opt: 1189, E(): 0, (52.8% identity in 337 aa overlap);

P21174|DP3B MICU from Micrococcus luteus (310 aa);  
 P52023|DP3B\_SYN7 from Synecococcus sp. strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in neighbouring cosmid MCY10H4.01."  
 /codon\_start=1  
 /evidence=experimental  
 /transl\_table=1  
 /product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA NUCLEOTIDYLTRANSFERASE)"  
 /protein\_id="CAD92864.1"  
 /db\_xref="GI:31616764"  
 /translation="MDAATRTVGLTDLTFLRLRESFADAVSWAKNLPARPAPVLSG  
 VLLTSGDNLGITISGDIYVSRAAQVAGIYSPGVLVSGRLSDITRALPNKPVGVHV  
 EGNRVALTGNARFSLFTMVEDYPTLPTPEETGLPFAELFAEASVIAARAGDDT  
 LPMLTGIRVEILGETVLAATDFRLAVLEKWSASPDIEAAVLVPATLAARAKAG  
 IGGSDRLSLGTGPGVKGKGLGIISNGKRSITRLLDAAFFKQLQLTETHTATMD  
 VAEILKILVALVADRGAGQVMEFADGKSVLSAGADLVGRAEEDLVVDYAGEPLTIA  
 FNYTITDGLSSLSRSERVSFGFTTAGKPAALLRPVSGDDRPVAGLNGSGFFPAVSTDYV  
 YILMPVRLPG"  
 3280..4437  
 /gene="recF"  
 /locus\_tag="Mb0003"  
 3280..4437  
 /gene="recF"  
 /locus\_tag="Mb0003"  
 /note="Mb0003, recF, len: 385 aa. Equivalent to Rv0003,  
 len: 385 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 385 aa overlap). recF, DNA replication and repair protein (see citations below), equivalent to others Mycobacterial DNA replication and repair proteins e.g. NP\_301131.1|NC 002677 from Mycobacterium leprae (385 aa); Q9L7L5|RECF MYCPA from Mycobacterium avium subsp. paratuberculosis (385 aa); P50916|RECF MYCSM from Mycobacterium smegmatis (384 aa); etc. Also highly similar to others e.g. P36176|RECF STRCO DNA REPLICATION AND REPAIR PROTEIN from Streptomyces coelicolor (373 aa); NP\_449892.1|NC 000911 from Synecocystis sp. strain PCC 6803 (384 aa); NP\_469352.1|NC\_003212 from Listeria innocua (370 aa); etc. Contains P500017 ATP/GTP-binding site motif A (p-loop), P500617 RecF protein signature 1, and P500618 RecF protein signature 2. BELONGS TO THE RECF FAMILY."  
 /codon\_start=1  
 /evidence=experimental  
 /transl\_table=1  
 /product="DNA REPLICATION AND REPAIR PROTEIN RECF (SINGLE-STRAND DNA BINDING PROTEIN)"  
 /protein\_id="CAD92865.1"  
 /db\_xref="GI:31616765"  
 /translation="MYVRLGLDRPSMACVDLELHPGTVFVGPNGYKTNLIALW  
 YSTLGHRSYADLPRIYGTDRAVISTTVNDGRECAVDLEIATGRVYKALNRSV  
 RSTRDVVLRVAVLPAPDLGLVGPDPADRRYLDLAIVRPAIAVRAEYVRVQ  
 RTALLKSVPGARYGRGVDTLVWDSLSAEGAEVLAARDLVNQLAPEVKAYQL  
 LAPESASIGYRASMDVTGPSEQSDTQQLAARLLAARARAEVGLVGPV  
 RDLILRLGDORAKGFASHEASLAVALLAAYQLLVDDGGEVLLDDVPAELDVM  
 RRALATPAESAEQVLTVAVLEDPAGWDARVHIDVADDTGSMVSLP"  
 4434..4997  
 /locus\_tag="Mb0004"  
 4434..4997  
 /locus\_tag="Mb0004"  
 /note="Mb0004, -, len: 187 aa. Equivalent to Rv0004, len: 187 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 187 aa overlap). Conserved hypothetical protein (see citation below), highly similar, but longer 21 aa in N-terminus, to AAF33696.1|AF222789 unknown protein from Mycobacterium avium subsp. paratuberculosis (166 aa); and highly similar to NP\_301132.1|NC\_002677 conserved hypothetical protein from Mycobacterium leprae (169 aa); S70990 hypothetical protein from Mycobacterium smegmatis (194 aa). Also similar to in C-terminus to C-terminal part of P35925|YREG\_STRCO HYPOTHETICAL 19, 8 KDA PROTEIN (IN RECF-GRB INTERGENIC REGION) from Streptomyces coelicolor (190 aa), FASTA scores: Opt: 404, E(): 3.9e-18, (40.7% identity in 189 aa overlap)."

gene  
/codon\_start=1  
/transl\_table=11  
/product="CONSERVED HYPOTHETICAL PROTEIN"  
/protein\_id="CAD92866.1"  
/db\_xref="GI:31616766"  
/translation="MTGSVDRPDQNGRERLMKSPGLDVRITLDEAARAARGODAG  
RGVAVASGRVARRRSGPDPDIRDPOLGKAARELAKRGWVRVAEGMVLQGW  
SAVGHQIAHASPALNDQVLVIAESTAWATQLRIMOQLAKIAAAGVNDVVRSL  
KITGPAPSRWKKGRHAGRPRTYG"  
5123..7267  
/gene="gyrB"  
/locus\_tag="Mb0005"  
5123..7267  
/locus\_tag="Mb0005"  
/gene="gyrB"  
/locus\_tag="Mb0005"  
/EC\_number="5.99.1.3"  
/note="Mb0005, gyrB, len: 714 aa. Equivalent to Rv0005,  
len: 714 aa, from Mycobacterium tuberculosis strain H37Rv,  
(99.9% identity in 714 aa overlap). gyrB, DNA gyrase  
subunit B (EC 5.99.1.3) (see citations below), equivalent,  
except in N-terminus, to other Mycobacterial DNA GYRASES  
SUBUNIT B e.g. T10005 from Mycobacterium leprae (697 aa);  
Q9L7L3 (GYR B MYCPA from Mycobacterium avium subsp.  
paratuberculosis (677 aa) (has its N-terminus shorter);  
P48355 (GYR B MYCSM from Mycobacterium smegmatis (675 aa);  
etc. Also highly similar to others e.g. T10969 from  
Streptomyces coelicolor (686 aa); P50075 (GYB5\_STRSH from

Alignment Scores:  
Pred. No.: 1 796-39 Length: 343050  
Score: 651.00 Matches: 128  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-684-215B-23 (1-128) x BX248334 (1-343050)

QY 1 ThrAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 152008 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 152067  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 152068 GGGCAGCGCATGCGCATCGCGGCCAGATCCGATCGGTGGGGGTGACCCACCGTTTCAAT 152127  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 152128 ATCGGCGCTACCGCTTCCTCGCTGGGTGTTGTCGACAAACGCGCAACGCGCGCAGCA 152187  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 152188 GTCCACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGCAGCTG 152247  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 152248 ATCACCGCGGTGACGCGCTCCGATCAACTCGGCCACCGCATGCGCGGACGCGTTAAC 152307  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 152308 GGGCATCATCCCGTACGCTCATCTCGTGACCTGCAACCAACCAAGTCGGCGCGCAGCGT 152367  
QY 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 152368 ACAGGAACTGACATTCGCCGAG 152391

RESULT 13  
AR169152  
LOCUS  
DEFINITION  
Sequence 4 from patent US 6290969.  
ACCESSION  
AR169152  
VERSION  
AR169152.1 GI:17906927  
KEYWORDS

gene  
/codon\_start=1  
/transl\_table=11  
/product="CONSERVED HYPOTHETICAL PROTEIN"  
/protein\_id="CAD92866.1"  
/db\_xref="GI:31616766"  
/translation="MTGSVDRPDQNGRERLMKSPGLDVRITLDEAARAARGODAG  
RGVAVASGRVARRRSGPDPDIRDPOLGKAARELAKRGWVRVAEGMVLQGW  
SAVGHQIAHASPALNDQVLVIAESTAWATQLRIMOQLAKIAAAGVNDVVRSL  
KITGPAPSRWKKGRHAGRPRTYG"  
5123..7267  
/gene="gyrB"  
/locus\_tag="Mb0005"  
5123..7267  
/locus\_tag="Mb0005"  
/gene="gyrB"  
/locus\_tag="Mb0005"  
/EC\_number="5.99.1.3"  
/note="Mb0005, gyrB, len: 714 aa. Equivalent to Rv0005,  
len: 714 aa, from Mycobacterium tuberculosis strain H37Rv,  
(99.9% identity in 714 aa overlap). gyrB, DNA gyrase  
subunit B (EC 5.99.1.3) (see citations below), equivalent,  
except in N-terminus, to other Mycobacterial DNA GYRASES  
SUBUNIT B e.g. T10005 from Mycobacterium leprae (697 aa);  
Q9L7L3 (GYR B MYCPA from Mycobacterium avium subsp.  
paratuberculosis (677 aa) (has its N-terminus shorter);  
P48355 (GYR B MYCSM from Mycobacterium smegmatis (675 aa);  
etc. Also highly similar to others e.g. T10969 from  
Streptomyces coelicolor (686 aa); P50075 (GYB5\_STRSH from

SOURCE  
ORGANISM  
Unknown.  
Unclassified.  
REFERENCE  
1 (bases 1 to 447)  
AUTHORS  
Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,  
Vedvick,T.S. and Twardzik,D.R.  
TITLE  
Compounds and methods for immunotherapy and diagnosis of  
tuberculosis  
JOURNAL  
Patent: US 6290969-A 4 18-SEP-2001;  
FEATURES  
Location/Qualifiers  
1..447  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 5 526-42 Length: 447  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservatives: 0  
Best Local Similarity: 99.22% Mismatches: 1  
Query Match: 99.22% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR169152 (1-447)

QY 1 ThrAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProfile 20  
Db 11 ACGGCGCGTCCGATTAACCTCCAGCTGTCACAGGTGGCGAGGATTCGCCATTCGGATC 70  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 71 GGGCAGCGCATGCGCATCGCGGCCAGATCCGATCGGTGGGGGTGACCCACCGTTTCAAT 130  
QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlyAsnGlyAlaArg 60  
Db 131 ATCGGCGCTACCGCTTCCTCGCTGGGTGTTGTCGACAAACGCGCAACGCGCGCAGCA 190  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 191 GTCCACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGCATCTCCACCGCGCAGCTG 250  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCGCGGTGACGCGCTCCGATCAACTCGGCCACCGCATGCGCGGACGCGTTAAC 310  
QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCCGTACGCTCATCTCGTGACCTGCAACCAACCAAGTCGGCGCGCAGCGT 370  
QY 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 371 ACAGGAACTGACATTCGCCGAG 394

RESULT 14  
AR182442  
LOCUS  
DEFINITION  
Sequence 4 from patent US 6338852.  
ACCESSION  
AR182442  
VERSION  
AR182442.1 GI:20225649  
KEYWORDS  
Unknown.  
ORGANISM  
Unknown.  
Unclassified.  
REFERENCE  
1 (bases 1 to 447)  
AUTHORS  
Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,  
Vedvick,T.S. and Twardzik,D.R.  
TITLE  
Compounds and methods for diagnosis of tuberculosis  
JOURNAL  
Patent: US 6338852-A 4 15-JAN-2002;  
FEATURES  
Location/Qualifiers  
1..447  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
 Pred. No.: 5,52e-42 Length: 447  
 Score: 646.00 Matches: 127  
 Percent Similarity: 99.22% Conservative: 0  
 Best Local Similarity: 99.22% Mismatches: 1  
 Query Match: 99.23% Indels: 0  
 Gaps: 0

US-09-684-215B-23 (1-128) x AR182442 (1-447)

QY 1 ThrAlaAlaSerAspAsnGlyGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
 Db 11 ACGCGCGGTCGATACCTCCAGCTCTCCAGGGTGGGCGAGGATTCGCCATTCGATC 70  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 40  
 Db 71 GGGCAGCGGATGCGGATCGGGCGGCGATCCGATCGGTTGGGGGTACCCACCGTTTCAT 130  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60  
 Db 131 ATCGGGCCCTACCGCTTCTCGGCTTGGGTGTGTGACAAACACGCGCAACGGCGCACGA 190  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 191 GTCCAAACGGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
 Db 251 ATCACCAGCGTTCGACGGCGCTCCGATCACTCGGCACCGCATCGCGACGCGCTTAAC 310  
 QY 101 GlyHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
 Db 311 GGGCATCATCCCGTTCAGCTCATCTCGGTGAACCTGGCAACCAAGTCGGGGCGACGGGT 370  
 QY 121 ThrGlyAsnValThrLeuAlaGlu 128  
 Db 371 ACAGGGAACGTGACATTCGCCGAG 394

RESULT 16  
 AR233097  
 LOCUS AR233097 447 bp DNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 4 from patent US 6458366.  
 ACCESSION AR233097  
 VERSION AR233097.1 GI:27275533  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS Read, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
 Vedvik, T.S., Twardzik, D.R., Lodes, M.J. and Hendrickson, R.C.  
 TITLE Compounds and methods for diagnosis of tuberculosis  
 JOURNAL Patent: US 6458366-A 4 01-OCT-2002;  
 FEATURES Location/Qualifiers  
 source  
 1..447  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,52e-42 Length: 447  
 Score: 646.00 Matches: 127  
 Percent Similarity: 99.22% Conservative: 0  
 Best Local Similarity: 99.22% Mismatches: 1  
 Query Match: 99.23% Indels: 0  
 Gaps: 0

US-09-684-215B-23 (1-128) x AR233097 (1-447)

QY 1 ThrAlaAlaSerAspAsnGlyGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
 Db 11 ACGCGCGGTCGATACCTCCAGCTCTCCAGGGTGGGCGAGGATTCGCCATTCGATC 70  
 QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 40  
 Db 71 GGGCAGCGGATGCGGATCGGGCGGCGATCCGATCGGTTGGGGGTACCCACCGTTTCAT 130  
 QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAlaAspAsnGlyValAlaArg 60  
 Db 131 ATCGGGCCCTACCGCTTCTCGGCTTGGGTGTGTGACAAACACGCGCAACGGCGCACGA 190  
 QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
 Db 191 GTCCAAACGGGTGTCGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250  
 QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

RESULT 15  
 AR194825  
 LOCUS AR194825 447 bp DNA linear PAT 20-APR-2002  
 DEFINITION Sequence 4 from patent US 6350456.  
 ACCESSION AR194825  
 VERSION AR194825.1 GI:20244262  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 447)  
 AUTHORS Read, S.G., Skeiky, Y.A.W. and Dillon, D.C.  
 TITLE Compositions and methods for the prevention and treatment of M.  
 tuberculosis infection  
 JOURNAL Patent: US 6350456-A 4 26-FEB-2002;  
 FEATURES Location/Qualifiers  
 source  
 1..447  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,52e-42 Length: 447  
 Score: 646.00 Matches: 127  
 Percent Similarity: 99.22% Conservative: 0  
 Best Local Similarity: 99.22% Mismatches: 1  
 Query Match: 99.23% Indels: 0  
 Gaps: 0

US-09-684-215B-23 (1-128) x AR194825 (1-447)

QY 1 ThrAlaAlaSerAspAsnGlyGlnLeuSerGlnGlyGlnGlyPheAlaIleProile 20  
 Db 11 ACGCGCGGTCGATACCTCCAGCTCTCCAGGGTGGGCGAGGATTCGCCATTCGATC 70

Db 251 ATCAGCGGTGCGAGCGCTCCGATCACTCCGACCGCGATGGCGGACGGCGTTAAC 310  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCCGGTGAGTCCATCTCGGTGAATGGCAACCAAGTCGGCGGCGACGCGT 370  
Qy 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 371 ACAGGGAACGTGACATTGGCCGAG 394  
RESULT 17  
AR353302  
LOCUS AR353302 447 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 4 from patent US 6592877.  
ACCESSION AR353302  
VERSION AR353302.1 GI:33759108  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S., Twardzik, D.R., Lodes, M.J. and Hendrickson, R.C.  
TITLE Compounds and methods for immunotherapy and diagnosis of  
tuberculosis  
JOURNAL Patent: US 6592877-A 4 15-JUL-2003;  
FEATURES Location/Qualifiers  
source 1. .447  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores: 5.52e-42 Length: 447  
Pred. No.: 646.00 Matches: 127  
Score: 99.22% Conservative: 0  
Percent Similarity: 99.22% Mismatches: 1  
Best Local Similarity: 99.23% Indels: 0  
Query Match: 99.23% Gaps: 0  
DB: 6  
US-09-684-215B-23 (1-128) x AR353302 (1-447)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 70  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTCAT 130  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 131 ATCGGGCTTACCGCTTCTCGGCTTGGGTGGTGTTCGACAAACACGCGGCGACGA 190  
Qy 61 ValGlnArgValValGlySerAlaProAlaIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 80  
Db 191 GTCCAAACGCGTGTGGGAGCGCTCCGCGGCGCATCTCCGCGCAAGTCTCCGCGGCGACGTG 250  
Qy 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCAGCGTGTGACGGCGCTCCGATCACTCCGCGCACCGCGATGGCGGCGCGCTTAAC 310  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCCGGTGAGTCCATCTCGGTGAATGGCAACCAAGTCGGCGGCGACGCGT 370  
Qy 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 371 ACAGGGAACGTGACATTGGCCGAG 394  
RESULT 18  
AX429596

LOCUS AX429596 447 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 4 from Patent EP1203817.  
ACCESSION AX429596  
VERSION AX429596.1 GI:21540845  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Reed, S.G., Skeiky, Y.A., Dillon, D.C., Campos-Neto, A., Houghton, R.L.,  
Vedvick, T.S. and Twardzik, D.R.  
TITLE Compounds and methods for immunotherapy and diagnosis of  
tuberculosis  
JOURNAL Patent: EP 1203817-A 4 08-MAY-2002;  
FEATURES Location/Qualifiers  
source 1. .447  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Alignment Scores: 5.52e-42 Length: 447  
Pred. No.: 646.00 Matches: 127  
Score: 99.22% Conservative: 0  
Percent Similarity: 99.22% Mismatches: 1  
Best Local Similarity: 99.23% Indels: 0  
Query Match: 99.23% Gaps: 0  
DB: 6  
US-09-684-215B-23 (1-128) x AX429596 (1-447)  
Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20  
Db 11 ACGGCCGCGTCCGATTAACCTCCAGCTGTCCAGGGTGGCGAGGATTCGCCATTCCGATC 70  
Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 71 GGGCAGCGCATGGCGATCGCGGCCAGATCCGATCGGGTGGGGGTCCACCCACCGTTCAT 130  
Qy 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60  
Db 131 ATCGGGCTTACCGCTTCTCGGCTTGGGTGGTGTTCGACAAACACGCGGCGACGA 190  
Qy 61 ValGlnArgValValGlySerAlaProAlaIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 80  
Db 191 GTCCAAACGCGTGTGGGAGCGCTCCGCGGCGCATCTCCGCGCAAGTCTCCGCGGCGACGTG 250  
Qy 81 IleThrAlaValAspGlyValAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 251 ATCACCAGCGTGTGACGGCGCTCCGATCACTCCGCGCACCGCGATGGCGGCGCGCTTAAC 310  
Qy 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120  
Db 311 GGGCATCATCCCGGTGAGTCCATCTCGGTGAATGGCAACCAAGTCGGCGGCGACGCGT 370  
Qy 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 371 ACAGGGAACGTGACATTGGCCGAG 394  
RESULT 19  
AX832581  
LOCUS AX832581 447 bp DNA linear PAT 12-DEC-2003  
DEFINITION Sequence 4 from Patent EP1347055.  
ACCESSION AX832581  
VERSION AX832581.1 GI:39840631  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Reed, S.G., Skeiky, Y.A., Campos-Neto, A., Houghton, R.L.,  
Vedvick, T.S., Twardzik, D.R. and Dillon, D.C.

TITLE Compounds for immunotherapy and diagnosis of tuberculosis

JOURNAL Patent: EP 1347055-A 4 24-SEP-2003;

source CORIXA CORPORATION (US)

FEATURES Location/Qualifiers

1..447

/organism="unidentified"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.: 5,52e-42 Length: 447

Score: 646.00 Matches: 127

Percent Similarity: 99.22% Conservatives: 0

Best Local Similarity: 99.22% Mismatches: 1

Query Match: 99.23% Indels: 0

DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AX832581 (1-447)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20

Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db 71 GGGCAGGCGATGGCGATCGGGCGCATCCGATCGGGTGGGGGTACCCACCGTTTCAT 130

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlnGlyAlaArg 60

Db 131 ATCGGGCTTACGGCTTCTCGGGTGGGTGTTGTCGACAAACGCGACGCGCAGCA 190

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 191 GTCCAAACGCTGTGGAGAGGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGAGCTG 250

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 251 ATCAACGGCGTCCGAGCGCTCCGATCACTCGGCCACCGCGATCGCGAGCGGCTTAAC 310

QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120

Db 311 GGGCATCATCCGGTCAAGTATCTCGGTGAACCTGGCAACCAAGTCCGGCGGCAACGCGT 370

QY 121 ThrGlyAsnValThrLeuAlaGlu 128

Db 371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 20

BD006325

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD006325 447 bp DNA linear PAT 31-JAN-2002

Compounds and methods for diagnosis of Tuberculosis.

BD006325

BD006325.1 GI:18634696

JP 2001500383-A/4.

unidentified

unidentified

unclassified.

1 (bases 1 to 447)

Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,

Vedvick, T.S., Twardzik, D.R., and Lodes, M.J.

Compounds and methods for diagnosis of Tuberculosis

Patent: JP 2001500383-A 4 16-JAN-2001;

CORIXA CORP

OS Unidentified

PN JP 2001500383-A/4

PD 16-JAN-2001

PP 07-OCT-1997 JP 1998518432

PR 11-OCT-1996 US 08/729622,13-MAR-1997 US 08/818111 PI

STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI ANTONIO CAMPOS

NETO,

PI RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI

MICHAEL J LODES

PC C12N15/31,C07K14/35,C07K16/12,C12Q1/68,C12N15/62,G01N33/53 CC

Strandedness: Single;

CC Topology: Linear;

PH Key Location/Qualifiers

FT source 1..447

/organism="Unidentified".

FEATURES Location/Qualifiers

1..447

/organism="unidentified"

/mol\_type="genomic DNA"

/db\_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.: 5,52e-42 Length: 447

Score: 646.00 Matches: 127

Percent Similarity: 99.22% Conservatives: 0

Best Local Similarity: 99.22% Mismatches: 1

Query Match: 99.23% Indels: 0

DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x BD006325 (1-447)

QY 1 ThrAlaAlaSerAspAsnGlnLeuSerGlnGlyGlnGlyGlnGlyPheAlaIleProIle 20

Db 11 ACGGCGCGTCCGATAACTTCCAGCTGTCCAGGGTGGGAGGATTCGGCATTCGGATC 70

QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db 71 GGGCAGGCGATGGCGATCGGGCGCATCCGATCGGGTGGGGGTACCCACCGTTTCAT 130

QY 41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnGlnGlyAlaArg 60

Db 131 ATCGGGCTTACGGCTTCTCGGGTGGGTGTTGTCGACAAACGCGACGCGCAGCA 190

QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 191 GTCCAAACGCTGTGGAGAGGCTCCGGCGCAAGTCTCGGCATCTCCACCGCGAGCTG 250

QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 251 ATCAACGGCGTCCGAGCGCTCCGATCACTCGGCCACCGCGATCGCGAGCGGCTTAAC 310

QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120

Db 311 GGGCATCATCCGGTCAAGTATCTCGGTGAACCTGGCAACCAAGTCCGGCGGCAACGCGT 370

BD006445 447 bp DNA linear PAT 31-JAN-2002

Compounds and methods for immunotherapy and diagnosis of

Tuberculosis.

BD006445

BD006445.1 GI:18634816

JP 2001501832-A/4.

unidentified

unidentified

unclassified.

1 (bases 1 to 447)

Reed, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,

Vedvick, T.S., Twardzik, D.R., and Lodes, M.J.

Compounds and methods for immunotherapy and diagnosis of

Patent: JP 2001501832-A 4 13-FEB-2001;

CORIXA CORP

OS Unidentified

PN JP 2001501832-A/4

PD 13-FEB-2001

PP 07-OCT-1997 JP 1998518456

PR 11-OCT-1996 US 08/730510,13-MAR-1997 US 08/818112 PI



Compounds and methods for immunotherapy and diagnosis of					
Patent: JP 2001517069-A 4	02-OCT-2001;				
CORIXA CORP					
OS Unidentified					
PN JP 2001517069-A/4					
PD 02-OCT-2001					
PP 30-AUG-1996	JP 1997511464				
PR 01-SEP-1995	US 08/523436, 22-SEP-1995	US	08/533634	PR	
PT 22-MAR-1996	US 08/620874, 05-JUN-1996	US	08/659683	PR	
12-JUL-1986	US 08/680574				
PI STEVEN G REED, YASIR A W SKEIKI, DAVIN C DILLON, ANTONIO CAMPOS					
PI NETO,					
PI RAYMOND HOUGHTON, THOMAS H VEDVICK, DANIEL R TWARDZIK PC					
C12N15/31, C07K3/4/35, A61K38/16, C12N15/62, G01N33/569, PC					
C12N5/10,					
C12N1/21//A61K39/04, (C12N1/21.C12R1:19)					
PC Strandedness: Single;					
CC Topology: Linear;					
CC Tuberculosis					
CC Compounds and methods for immunotherapy and diagnosis of CC					

五五

```

1. .447
/organism='Unidentified'
Location/Qualifiers
1. .447
/organism="unidentified"
/mol_type="Genomic DNA"
/ab_xref="taxon:32544"

```

## ORIGIN

[illegible]

and

101	GLYHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg	120
102		
103		
104		
105		
106		
107		
108		
109		
110		
111	GGGCAATCATCCCGGTGACGCATCTCGTGAACGGCAACAGTCGGGCGGCACGCT	129
112		
113		
114		
115		
116		
117		
118		
119		
120	ThrGlyAsnValThrLeuAlaGlu	128
121		
122		
123		
124		
125		
126		
127		
128		
129	ACAGGGAACTGCATTTGGCCGAG	394
130		
131		
132		
133		
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		

Qy

111

DB

RECEIVED 23

RESULT 23  
BD205817

LOCUS

### DEFINITION

ACCESSION



VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BD205817.1 GI:33015587  
JP 2002530050-A/4  
Mycobacterium tuberculosis  
Mycobacterium tuberculosis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.  
1 (bases 1 to 447)  
Read, S.G., Skeiky, Y.A.W., Dillon, D.C., Neto, A.C., Houghton, R.,  
Vedvick, T.S., Twardzik, D.R., Lodes, M.J. and Hendrickson, R.C.  
Compounds and methods for diagnosis of tuberculosis  
Patent: JP 2002530050-A 4 17-SEP-2002;  
CORIXA CORP

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

OS Mycobacterium tuberculosis  
PN JP 2002530050-A/4  
PD 17-SEP-2002  
PP 17-FEB-1999 JP 2000531132  
PR 18-FEB-1998 US 09/024753 05-MAY-1998 US 05/072596 PI  
STEVEN G REED, YASIR A W SKEIKY, DAVIN C DILLON, ANTONIO CAMPOS PI  
NETO,  
RAYMOND HOUGHTON, THOMAS S VEDVICK, DANIEL R TWARDZIK, MICHAEL J  
LODES,  
RONALD C HENDRICKSON  
PC C12N15/09, C07H21/02, C07K14/35, C07K16/12, PC  
C07K17/00,  
PC C07K19/00, C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/04, C12Q1/  
PC 68,  
PC G01N33/569, G01N33/68//A61K38/00, A61K39/04, A61K39/395, A61K39/  
PC 395, A61P31/06,  
PC (C12N1/21, C12R1/19), C12N5/00, C12N5/00, A61K37/02 CC  
Compounds and methods for diagnosis of tuberculosis. FH Key  
FT source i. .447  
FT Location/Qualifiers  
/organism="Mycobacterium tuberculosis".  
1. .447  
Location/Qualifiers  
/organism="Mycobacterium tuberculosis"

ORIGIN

Alignment Scores:  
Pred. No.: 5,52e-42 Length: 447  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservative: 0  
Best Local Similarity: 99.22% Mismatches: 1  
Query Match: 99.23% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x BD205817 (1-447)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

Db 11 ACGGCCGCTCCGATTAACCTCCAGCTGTCAGGTCGGCAGGGATTCGCCATTCGGATC 70

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db 71 GGGCAGCGCATGCGATCGGGGCCAGATCCGATCCGATCGGGTGGGGGTCACCCACCGTTTCAT 130

Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValAsnGlyAlaArg 60

Db 131 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTGTCACACACGCGACCGCGCAGCA 190

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 191 GTCCACCGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCAGCTG 250

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 251 ATCACCGCGTCCGACGCGCTCCGATCACTCGGCACCGGATGGGACCGGCTTAAC 310

Qy 101 GlyHisHisProGlyAspValIleSerValThrTriPdnThrIlySerGlyGlyThrArg 120

Qy 121 ThrGlyAsnValThrLeuAlaGlu 128

Db 1118 ACAGGACGTCGATTTGGCCGAG 1141

RESULT 25

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

AR182455

Sequence 17 from patent US 6338852.

AR182455

AR182455.1 GI:20225662

Db 311 GGGCATCATCCGGTGACGTGACGTATCTGGTGAACTGGGCAACAGTCGGCGGCACGCGT 370

Qy 121 ThrGlyAsnValThrLeuAlaGlu 128

Db 371 ACAGGACGTCGATTTGGCCGAG 394

RESULT 24

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

Unknown.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .1872

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 2,33e-41 Length: 1872  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservative: 0  
Best Local Similarity: 99.22% Mismatches: 1  
Query Match: 99.23% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR169165 (1-1872)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlnGlyPheAlaIleProfile 20

Db 758 ACGGCCGCTCCGATTAACCTCCAGCTGTCAGGTCGGCAGGGATTCGCCATTCGGATC 817

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db 818 GGGCAGCGCATGCGATCGGGGCCAAATCCGATCCGATCGGGTGGGGGTCACCCACCGTTTCAT 877

Qy 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValAsnGlyAlaArg 60

Db 878 ATCGGGCTACCGCTTCCTCGGCTTGGGTGTGTCACACACGCGCACCGCGCAGCA 937

Qy 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db 938 GTCCACCGCTGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCAGCTG 997

Qy 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db 998 ATCACCGCGTCCGACGCGCTCCGATCACTCGGCACCGCGATGGCGGACCGGCTTAAC 1057

Qy 101 GlyHisHisProGlyAspValIleSerValThrTriPdnThrIlySerGlyGlyThrArg 120

Db 1058 GGGCATCATCCGGTGACGTGATCTCGTGAACTGGGCAACCAAGTCGGCGGCACGCGT 1117

Qy 121 ThrGlyAsnValThrLeuAlaGlu 128

Db 1118 ACAGGACGTCGATTTGGCCGAG 1141

RESULT 25

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

AR182455

Sequence 17 from patent US 6338852.

AR182455

AR182455.1 GI:20225662

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1872)  
AUTHORS Read, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S., and Twardzik, D.R.  
TITLE Compounds and methods for diagnosis of tuberculosis  
JOURNAL Patent: US 6338852-A 17 15-JAN-2002;  
FEATURES Location/Qualifiers  
source 1..1872  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 2,33e-41 Length: 1872  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservative: 0  
Best Local Similarity: 99.22% Mismatches: 1  
Query Match: 99.23% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR182455 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGGCGCGTCCGATAACTTCCAGCTGTCACAGGTCGGCGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 818 GGGCAGCGATGGCGATCCGGGCGCAATCCGATCGGTGGGGGTCAACCACCGTTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValAlaArg 60  
Db 878 ATCGGCGCTACCGCTTCTCGGCTGGGTGTGTGCGACAAACACGGCGGCACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCAAACGGTGTGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCGATG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCACCAGCGTTCGACGGCGTCCGATCACTCGGCACCGCGATCGCGACGCTTAAC 1057  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 1058 GGGCATCATCCCGTGGTGCATCTCGGTGAACCTGGCAACCAAGTCGGCGGCACGCGT 1117  
QY 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 26  
LOCUS AR194838 1872 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 17 from patent US 6350456.  
ACCESSION AR194838  
VERSION AR194838.1 GI:20244275  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1872)  
AUTHORS Read, S.G., Skeiky, Y.A.W., and Dillon, D.C.  
TITLE Tuberculosis infection  
JOURNAL Patent: US 6350456-A 17 26-FEB-2002;  
FEATURES Location/Qualifiers  
source 1..1872  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 2,33e-41 Length: 1872  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservative: 0  
Best Local Similarity: 99.23% Mismatches: 1  
Query Match: 99.23% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR194838 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGGCGCGTCCGATAACTTCCAGCTGTCACAGGTCGGCGAGGATTCGCCATTCGGATC 817  
QY 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40  
Db 818 GGGCAGCGATGGCGATCCGGGCGCAATCCGATCGGTGGGGGTCAACCACCGTTTCAT 877  
QY 41 IleGlyProThrAlaPheLeuGlyValValAspAsnGlyValAlaArg 60  
Db 878 ATCGGCGCTACCGCTTCTCGGCTGGGTGTGTGCGACAAACACGGCGGCACGA 937  
QY 61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80  
Db 938 GTCCAAACGGTGTGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGCGCGATG 997  
QY 81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100  
Db 998 ATCACCAGCGTTCGACGGCGTCCGATCACTCGGCACCGCGATCGCGACGCTTAAC 1057  
QY 101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg 120  
Db 1058 GGGCATCATCCCGTGGTGCATCTCGGTGAACCTGGCAACCAAGTCGGCGGCACGCGT 1117  
QY 121 ThrGlyAsnValThrLeuAlaGlu 128  
Db 1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 27  
LOCUS AR233110 1872 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 17 from patent US 6458366.  
ACCESSION AR233110  
VERSION AR233110.1 GI:27275546  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1872)  
AUTHORS Read, S.G., Skeiky, Y.A.W., Dillon, D.C., Campos-Neto, A., Houghton, R.,  
Vedvick, T.S., Twardzik, D.R., Lodes, M.J., and Hendrickson, R.C.  
TITLE Compounds and methods for diagnosis of tuberculosis  
JOURNAL Patent: US 6458366-A 17 01-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..1872  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 2,33e-41 Length: 1872  
Score: 646.00 Matches: 127  
Percent Similarity: 99.22% Conservative: 0  
Best Local Similarity: 99.22% Mismatches: 1  
Query Match: 99.23% Indels: 0  
DB: 6 Gaps: 0

US-09-684-215B-23 (1-128) x AR233110 (1-1872)

QY 1 ThrAlaAlaSerAspAsnPhenGlnLeuSerGlnGlyGlnGlyPheAlaIleProIle 20  
Db 758 ACGGCGCGTCCGATAACTTCCAGCTGTCACAGGTCGGCGAGGATTCGCCATTCGGATC 817

Qy	21	GlycAlaMetAlaIleAlaGlyGlnIleAArgSerGlyGlyGlySerProThrValHis	40					
Db	818	GGG CAG CGC GAT GGC GAT CCG GGG CCA AAT CCG AT CCG GTG GGG GGT CAC CCA CCG GTT CAT	877					
Qy	41	IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg	60					
Db	878	AT CCG GCT ACC GCT TCT CCG CTT GGG TGT TGT CGA CAA CAA CGC ACG CCG CAC GCA	937					
Qy	61	ValGlnArgValValGlySerAlaProAlaAAserLeuGlyIleSerThrGlyAspVal	80					
Db	938	GT CCA ACG CGT GGT CGG AAG CGT CCG CGC GCA GT CCG CCA TCT CCA CCG CGC ACG GT	997					
Qy	81	IleThrAlaValAspGlyAlaProIleAAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100					
Db	998	AT CAC CGC GGT GAC CGC GGT CCG AT CAA CTC GGC CAC CGC GAT GGC GAC GCG CTT AAC	1057					
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIlySerGlyGlyThrArg	120					
Db	1058	GGG CAT CAT CCG GT GAC GT CAT CTC GGT GAA CTC GCA AACT GGC AACA AAG TCG GCG CAC GCG T	1117					
Qy	121	ThrGlyAsnValThrLeuAlaGlu	128					
Db	1118	AC AGG AAC GTG AC ATT GGC CG AG	1141					
RESULT 28								
AR353315								
LOCUS	AR353315	1872 bp	DNA					
DEFINITION	Sequence 17 from patent US 6592877.							
ACCESSION	AR353315							
VERSION	AR353315.1	GI:33759121						
KEYWORDS								
SOURCE	Unknown.							
ORGANISM	Unknown.							
REFERENCE	Unclassified.							
AUTHORS	1 (Bases 1 to 1872)							
TITLE								
Patent: US 6592877-A 17 15-JUL-2003;								
Location/Qualifiers								
1..1872								
/organism="unknown"								
/mol_type="genomic DNA"								
ORIGIN								
Alignment Scores:								
Pred. No.:	2,33e-41	Length:	1872					
Score:	646.00	Matches:	127					
Percent Similarity:	99.22%	Conservative:	0					
Best Local Similarity:	99.22%	Mismatches:	1					
Query Match:	99.23%	Indels:	0					
DB:	6	Gaps:	0					
US-09-684-215B-23 (1-128) x AR353315 (1-1872)								
Qy	1	ThrAlaAAserAspAsnPheGlnIleuSerGlnGlyGlnGlyPheAlaIleProIle	20					
Db	758	ACG GCG CGC GTC CGA TAA CTT CCA GCT GTC CCA GGG TGG CAG GGA TTC GCC ATT CCG ATC	817					
Qy	21	GlyGlnAAserAlaIleAlaGlyGlnIleAArgSerGlyGlyGlySerProThrValHis	40					
Db	818	GGC AGC GCA TCG CCA TCG GGC CCA AAT CCA GAT CCG GTG GGG GGT CAC CCA CCG GTT CAT	877					
Qy	41	IleGlyProThrAlaPheLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg	60					
Db	878	AT CCG GCT ACC GCT TCT CCG CTT GGG TGT TGT CGA CAA CAA CGC ACG CCG CAC GCA	937					
Qy	61	ValGlnArgValValGlySerAlaProAlaAAserLeuGlyIleSerThrGlyAspVal	80					
Db	938	GT CCA ACG CGT GGT CGG AAG CGT CCG CGC GCA GT CCG CCA TCT CCA CCG CGC ACG GT	997					
Qy	81	IleThrAlaValAspGlyAlaProIleAAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100					

Db	998	ATCACCGCGTGCAGCGCGCTCCGATCAACTCGGCCACCGCATGCGGCGCTTAAC	1057
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg	120
Db	1058	GGGCATCATCCCGGTGACGTCACTCGGTGAACCTGGCAACCAAGTCGGCGGCACGCGT	1117
Qy	121	ThrGlyAsnValThrLeuAlaGlu	128
Db	1118	ACAGGGAACGTGACATTGGCCGAG	1141
RESULT 29			
LOCUS	AX429609	1872 bp	DNA
DEFINITION	Sequence 17 from Patent EP1203817.		linear
ACCESSION	AX429609		
VERSION	AX429609.1	GI:21540858	
KEYWORDS			
SOURCE	unidentified		
ORGANISM	unidentified		
REFERENCE	unclassified.		
AUTHORS	1. Reed, S.G., Skeiky, Y.A., Dillon, D.C., Campos-Neto, A., Houghton, R.L.,		
TITLE	Vedvick, T.S. and Twardzik, D.R.		
JOURNAL	Compounds and methods for immunotherapy and diagnosis of		
FEATURES	tuberculosis		
source	Patent: EP 1203817-A 17 08-MAY-2002;		
	CORIXA CORPORATION (US)		
ORIGIN	Location/Qualifiers		
	1..1872		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
Alignment Scores:			
Pred. No.:	2.33e-41	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	6	Gaps:	0
US-09-684-215B-23 (1-128) x AX429609 (1-1872)			
Qy	1	ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle	20
Db	758	ACGCGCGGTCGATTAATCTCCAGTGTCCAGGTGGCAGGATTCGCCATTCGATC	817
Qy	21	GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis	40
Db	818	GGCGAGCGCGATGGCGATCGCGGGCCAAATCCGATCGGTGGGGGTCAACCCACGCTTCA	877
Qy	41	IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg	60
Db	878	ATCGGCGCTACCGCTTCTCCGCTTGGGTGTTCGACAAACAACGGCAACGGCGACGA	937
Qy	61	ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal	80
Db	938	GTCCAACCGGTGGTGGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACGGCGACGTG	997
Qy	81	IleThrAlaValAspGlyValAProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn	100
Db	998	ATCACCGCGGTGCAGCGCGCTCCGATCAACTCGGCCACCGCGATGGCGAGCGCTTAAC	1057
Qy	101	GlyHisHisProGlyAspValIleSerValThrTrpGlnThrIleSerGlyGlyThrArg	120
Db	1058	GGGCATCATCCCGGTGACGTCACTCGGTGAACCTGGCAACCAAGTCGGCGGCACGCGT	1117
Qy	121	ThrGlyAsnValThrLeuAlaGlu	128
Db	1118	ACAGGGAACGTGACATTGGCCGAG	1141